In the format provided by the authors and unedited.

nature

chemistry

# Atomically precise organomimetic cluster nanomolecules assembled via perfluoroaryl-thiol S<sub>N</sub>Ar chemistry

Elaine A. Qian<sup>1,2,4</sup>, Alex I. Wixtrom<sup>1</sup>, Jonathan C. Axtell<sup>1</sup>, Azin Saebi<sup>1</sup>, Dahee Jung<sup>1,4</sup>, Pavel

Rehak<sup>3</sup>, Yanxiao Han<sup>3</sup>, Elamar Hakim Moully<sup>1</sup>, Daniel Mosallaei<sup>1</sup>, Sylvia Chow<sup>1</sup>, Marco S.

Messina<sup>1,4</sup>, Jing Yang Wang<sup>1</sup>, A. Timothy Royappa<sup>5</sup>, Arnold L. Rheingold<sup>6</sup>, Heather D.

Maynard<sup>1,4</sup>, Petr Král<sup>3,7,8</sup> and Alexander M. Spokoyny<sup>1,4</sup>\*

<sup>1</sup>Department of Chemistry and Biochemistry, University of California, Los Angeles, 607 Charles E. Young Drive East, Los Angeles, California 90095, USA.

<sup>2</sup>Department of Bioengineering, University of California, Los Angeles, 410 Westwood Plaza, Los Angeles, California 90095, USA.

<sup>3</sup>Department of Chemistry, University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607, USA.

<sup>4</sup>California NanoSystems Institute, University of California, Los Angeles, 570 Westwood Plaza, Los Angeles, California 90095, USA.

<sup>5</sup>Department of Chemistry, University of West Florida, 11000 University Parkway, Pensacola, Florida 32514, USA.

<sup>6</sup>Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA.

<sup>7</sup>Department of Physics, University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607, USA.

<sup>8</sup>Department of Biopharmaceutical Sciences, University of Illinois at Chicago, 833 South Wood Street, Chicago, Illinois 60612, USA.

\*Corresponding author: Alexander M. Spokoyny Email: <u>spokoyny@chem.ucla.edu</u> Phone: (+1) 310-357-1327

# **Table of Contents**

General considerations		
Materials		
Instruments	5	
2D diffusion-ordered (DOSY) <sup>1</sup> H NMR spectroscopy	6	
High resolution transmission electron microscopy (HRTEM)	6	
Size exclusion chromatography-multi angle light scattering (SEC-MALS)	6	
Surface plasmon resonance (SPR)		
X-ray data collection and processing parameters		
Microwave synthesis	8	
Synthesis of 1	9	
Synthesis of 2	9	
Synthesis of 4-pentafluorophenyl(hydroxymethyl) benzene	10	
Synthesis of 4-pentafluorophenyl(bromomethyl) benzene		
Synthesis of <b>3</b>		
Synthesis of <b>2a</b>		
Synthesis of <b>2b</b>	14	
Synthesis of <b>2c</b>		
Synthesis of <b>2d</b>		
Synthesis of <b>2e</b>		
Synthesis of <b>2f</b>		
Synthesis of <b>2i</b>		
Synthesis of mPEGthiol <sub>766</sub> [ <b>j</b> ]		

Synthesis of <b>2</b> j	22
Synthesis of <b>2k</b>	23
Synthesis of <b>2</b> l	24
Synthesis of <b>3a</b>	25
Synthesis of <b>3b</b>	26
Synthesis of <b>3c</b>	27
Synthesis of <b>3d</b>	28
Synthesis of <b>3e</b>	29
Synthesis of <b>3f</b>	30
Synthesis of <b>3g</b>	31
Synthesis of <b>3h</b>	32
Synthesis of <b>3i</b>	33
Synthesis of <b>3</b> j	34
Synthesis of <b>3k</b>	35
Synthesis of <b>3</b> l	36
NMR and mass spectra	38
2D DOSY <sup>1</sup> H NMR spectroscopy plots of <b>2i-k</b> and <b>3i-k</b>	198
Stability studies of <b>2i</b> under biologically relevant conditions	204
Plot of conjugation progress of Boc-cysteine (g) onto clusters $2/3$	222
Computational work	223
References	238

# **Experimental Section**

**General considerations.** Microwave synthesis reactions and all post-microwave work-up and characterization were performed under ambient conditions. For the purposes of this manuscript, "ambient conditions" refer to room temperature (20 - 28 °C) and uncontrolled laboratory air.

Materials. Deuterated solvents were purchased from Cambridge Isotope Laboratories. MilliQ water described in this manuscript refers to purified potable water with a resistivity at 25 °C of  $\leq 18.2 \text{ M}\Omega \cdot \text{cm}$ . [NEt<sub>3</sub>H]<sub>2</sub>[B<sub>12</sub>H<sub>12</sub>] was purchased from Boron Specialties. EtOH (200 proof) was purchased from Decon Labs. Fmoc-L-amino acids (>98.5%) were purchased from Chem-Impex International, Inc. Piperidine (99%) was purchased from Spectrum. CaCl<sub>2</sub>·2 H<sub>2</sub>O (≥99%), MgCl<sub>2</sub>·6  $H_2O$  ( $\geq 99\%$ ), MnCl<sub>2</sub>·4  $H_2O$  ( $\geq 98\%$ ), diethyl ether (anhydrous,  $\geq 99.9\%$ ), glycine (98%), and Gibco minimum essential medium were purchased from Fisher Scientific. Thiophenol (99%) and poly(ethylene glycol) methyl ether (average MW 750 Da, MW range 715 - 785 Da) were purchased from Acros Organics. HBS-P pH 7.4 buffer (10 mM HEPES, 0.005% v/v Tween P20) and 1 M ethanolamine HCl (pH 8.5) were purchased from GE Healthcare Life Sciences. Fetal bovine serum was purchased from Sciencell Research Laboratories. FeCl<sub>3</sub>·6 H<sub>2</sub>O (≥97%), CsOH·1 H<sub>2</sub>O (≥99.5%), H<sub>2</sub>O<sub>2</sub> (30% in H<sub>2</sub>O), [N<sup>n</sup>Bu<sub>4</sub>]OH (40% in H<sub>2</sub>O), trifluoroacetic acid (TFA, 99%), triisopropylsilane (98%), N,N-dimethylformamide (DMF, ≥99.8%; anhydrous, 99.8%), MeCN (≥99.9%), CH<sub>2</sub>Cl<sub>2</sub> (≥99.5%), ethyl acetate (≥99.5%), hexanes (≥98.5%), MeOH (≥99.8%), N,Ndiisopropylethylamine (≥99%), tetrabutylammonium hexafluorophosphate (>99.0%, electrochemical grade), 1,2-ethanedithiol (≥98%), 1-hexanethiol (95%), benzyl mercaptan (99%), cysteamine (95%), 2-mercaptoethanol (≥99%), 1-thioglycerol (≥97%), O-(2-mercaptoethyl)-O'methyl-hexa(ethylene glycol) (average  $M_n$  356.48 Da,  $\geq$ 95%), O-(2-mercaptoethyl)-O'methylpolyethylene glycol (average M<sub>w</sub> 2,000 Da), 1-thio-β-D-glucose tetraacetate (97%), N-

(*tert*-Butoxycarbonyl)-L-cysteine methyl ester (97%), isopropoxytrimethylsilane (98%), K<sub>3</sub>PO<sub>4</sub> ( $\geq$ 98%), K<sub>2</sub>CO<sub>3</sub> ( $\geq$ 99%), Tris ( $\geq$ 99%), and triethylamine ( $\geq$ 99%) were purchased from Sigma-Aldrich. All reagents were used as received unless otherwise indicated.

Instruments. Bruker AV300, AV400, AV500, and DRX500 spectrometers were used to obtain <sup>1</sup>H, <sup>11</sup>B, <sup>13</sup>C{<sup>1</sup>H}, and <sup>19</sup>F NMR spectra and Bruker Topspin software was used to process the NMR data. <sup>1</sup>H and <sup>13</sup>C{<sup>1</sup>H} NMR spectra were referenced to residual solvent resonances in deuterated solvents (due to high humidity, H<sub>2</sub>O resonances are often present). <sup>11</sup>B and <sup>19</sup>F NMR spectra were referenced to BF<sub>3</sub>·Et<sub>2</sub>O and CFCl<sub>3</sub> external standards, respectively, at  $\delta$  0.0. *in situ* <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy was run unlocked and unshimmed. <sup>11</sup>B NMR spectra were baseline-corrected using the cubic spline correction tool within the Bruker Topspin software. Mass spectrometry data were acquired using a Thermo Scientific Q-Exactive Plus instrument with a quadrupole mass filter and Orbitrap mass analyzer or a Waters LCT Premier TOF system with ACQUITY LC and autosampler. IR spectroscopy was acquired on solid samples using a PerkinElmer Spectrum Two FT-IR spectrometer equipped with a diamond universal ATR probe. High resolution transmission electron microscopy (HRTEM) images were acquired with a FEI Titan electron microscope operating at 300 kV. Size exclusion chromatography-multi angle light scattering (SEC-MALS) was conducted on a GE AKTA PURE chromatographic system equipped with a WYATT miniDawn Treos MALS, WYATT optilab T-rEX RI detector, one Tosoh PWXL guard column (6.0 mm ID x 4.0 cm, 12 µm), and one Tosoh G4000PWxl (7.8 mm ID x 30 cm, 10 µm) column. Surface plasmon resonance (SPR) experiments were run on a GE Healthcare Life Sciences Biacore T100 instrument. Purification of peptides was done using a Waters HPLC system equipped with a UV/Vis detector set at  $\lambda = 214$  nm.

**2D diffusion-ordered (DOSY)** <sup>1</sup>**H NMR spectroscopy.** 2D DOSY experiments on purified samples of PEGylated OCNs were performed in D<sub>2</sub>O at 30 °C on a Bruker AV 300 spectrometer. The data were processed with the standard Bruker Topspin software – the T1/T2 *vargrad* fitting function was used to determine the diffusion coefficients. 2D DOSY plots were created with the Bruker Topspin software. Hydrodynamic diameters were estimated based on the diffusion coefficients using the Stokes-Einstein Equation.

**High resolution transmission electron microscopy (HRTEM).** HRTEM samples were prepared by dropping 5  $\mu$ L of 25  $\mu$ g/mL aqueous sample solutions onto carbon copper grids (Ted Pella). The samples were then blotted once with a filter paper and then left to air-dry for 10 minutes. Then, 3  $\mu$ L of a 2% w/w uranyl acetate aqueous solution was dropped on the grids, and subsequently blotted after 2 minutes.

Size exclusion chromatography-multi angle light scattering (SEC-MALS). Samples for SEC-MALS were prepared by dissolving sample in MilliQ water and filtering sample through a 0.20  $\mu$ m PTFE Fisherbrand syringe filter. Eluent was Millipore filtered MilliQ water with 0.02% NaN<sub>3</sub> at 12 °C (flow rate: 0.70 mL/min). Chromatograms were analyzed using Astra 6.0 software.

**Surface plasmon resonance (SPR).** All experiments were performed on a Biacore T100 instrument with a Series S CM5 chip (GE Healthcare Life Sciences). The procedure used here was modified from a published procedure by Safina *et al.*<sup>1</sup> The running buffer was 10 mM HEPES buffer (pH 7.4) with 0.005% Tween P20, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, and 1 mM MnCl<sub>2</sub>. 5  $\mu$ L/min flow rate was used throughout the experiments. First, a reference channel (flow cell 1) was prepared by activating the surface with a 0.4 M EDC and 0.1 M NHS (1:1 v/v) mixture during 30 minutes, then 1 M ethanolamine HCl (pH 8.5) during 40 minutes. Then, the sample channel (flow cell 2) was activated using under the EDC/NHS conditions, followed by injection of 0.1 mg/mL

ConA for 40 minutes and then 1 M ethanolamine HCl for 30 minutes for blocking. Analyte samples of 0.022  $\mu$ M to 130  $\mu$ M were injected in tandem over both cells for 6 minutes. Surfaces were regenerated by injecting 10 mM HCl for 2 minutes followed by injecting 10 mM glycine HCl (pH 2.5) for 2 minutes. Binding curves at various analyte concentrations were fitted to the Langmuir 1:1 binding model for an estimation of the binding constants. For the purpose of figure presentation, the sensorgrams were processed using the smoothing function in the OriginPro data analysis software.

**X-ray data collection and processing parameters.** For **2**, a single crystal was mounted on a nylon loop using perfluoropolyether oil and cooled rapidly to 100 K with a stream of cold dinitrogen. Diffraction data were measured using a Bruker APEX-II CCD diffractometer using Mo- $K_{\alpha}$  radiation. The cell refinement and data reduction were carried out using Bruker SAINT and the structure was solved with SHELXS-97. All subsequent crystallographic calculations were performed using SHELXL-2013. For **3**, single-crystal diffraction data were collected at 100(2) K on a Bruker Apex II CCD diffractometer with Mo K $\alpha$  radiation ( $\lambda = 0.71073$  Å). After correcting for absorption and polarization effects, structure solution and refinement were carried out using the SHELXT<sup>2</sup>, XL<sup>3</sup> and Olex2<sup>4</sup> software suites. Non-hydrogen atoms were refined with anisotropic thermal displacement parameters, and hydrogen atoms were placed in suitable riding positions.





7

compound		3	
empirical formula		$C_{78}H_{36}B_6F_{30}O_6$	
fw		1703.95	
temp / K		100	
wavelength / Å		0.71073 Å	
space group		P -1	
a / Å		19.211(3)	
b / Å		19.674(3)	
c / Å		22.866(4)	
$\alpha / \deg$		97.606(5)	
$\beta$ / deg		114.089(5)	
$\gamma / \deg$		109.756(5)	
V / Å		7047.9(18)	
Ζ		2	
d (calcd) / Mg·m <sup>-3</sup>		1.606	
abs coeff / mm <sup>-1</sup>		0.153	
R indices:	$R_1 =$	0.1771	
	$R_w =$	0.2030	

**Microwave synthesis.** Microwave reactions were performed using a CEM Discover SP microwave synthesis reactor. Except where noted otherwise, all reactions were performed in glass 10 mL microwave reactor vials purchased from CEM with silicone/PTFE caps. Flea micro PTFE-coated stir bars (10 mm x 3 mm) were used in the vials with magnetic stirring set to high and 15 seconds of premixing prior to the temperature ramping. All microwave reactions were carried out at 140 °C with the pressure release limit set to 250 psi (no reactions exceeded this limit to trigger venting) and the maximum wattage set to 250 W (the power applied was dynamically controlled by the microwave instrument and did not exceed this limit for any reactions). Column chromatography was performed using 2.0 - 2.25 cm inner diameter glass fritted chromatography columns with 20-30 cm of slurry-packed silica gel to ensure full separation of reagents and products. Unfiltered pressurized air was used to assist column chromatography.

#### Synthesis of 1



The  $[N^nBu_4]_2$  salt of  $[B_{12}(OH)_{12}]^{2-}$  was prepared according to the procedures detailed in Wixtrom *et al.* 2016.<sup>5</sup> *From this point,*  $N^nBu_4$  *will be referred to as TBA. Note:* **1** *is air-stable, but hygroscopic. Store under inert atmosphere or in a sealed desiccator to prevent excess absorption of water over extended periods of time in storage.* 

#### Synthesis of 2



Previously reported protocol<sup>6</sup> used to synthesize compound **2** – procedure is duplicated here. Compound **1** (300 mg, 0.366 mmol) was transferred out of a nitrogen filled glovebox, opened to the air, and dissolved in 4 mL acetonitrile in a 30 mL glass microwave vial. *N*,*N*diisopropylethylamine (1.21 mL, 6.96 mmol) and 2,3,4,5,6-pentafluorobenzyl bromide (6.86 mL 45.4 mmol) were added along with a magnetic stir bar, the vial was sealed with a Teflon/silicone cap, and the reaction mixture was heated under microwave conditions at 140°C with high stirring for 15 minutes. The volatiles were removed via rotary evaporation, and the excess reagent was eluted through a silica column with 65/35 hexanes/ethyl acetate, and the pink/purple product mixture was eluted with acetone. The acetone was removed via rotary evaporation and the residue was dissolved in ~5 mL 90/5/5 ethanol/acetonitrile/H<sub>2</sub>O. FeCl<sub>3</sub>·6H<sub>2</sub>O (1.88 g, 6.96 mmol) was added and the mixture was left to stir for 24 hours. The mixture was concentrated *in vacuo*. The residue (while still in the round bottom flask) was rinsed three times with water. The residue was then taken up in toluene and extracted three times with water. The organic fractions were combined and dried under vacuum. The resulting solid was charged with hexane and isolated by filtration to afford an orange/yellow solid (574 mg, 63%). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.23 (s, 24H). <sup>11</sup>B NMR (160 MHz, CDCl<sub>3</sub>):  $\delta$  40.9. <sup>13</sup>C{<sup>1</sup>H} NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  60.1. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>):  $\delta$  -145.1 (d, 24F, *-ortho*), -152.2 (t, 12F, *-para*), -161.3 – -161.5 (m, 24F, *- meta*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>84</sub>H<sub>84</sub>B<sub>12</sub>O<sub>12</sub> (M<sup>-</sup>), 2494.1499 Da; found, 2494.1631 Da. Crystallized from CDCl<sub>3</sub> at room temperature for 1 week to obtain a single crystal for X-ray diffraction analysis.

#### Synthesis of 4-pentafluorophenyl(hydroxymethyl) benzene



A solution of 4-pentafluorophenyl benzaldehyde (0.900 g, 3.30 mmol) and sodium borohydride (0.150 g, 3.96 mmol) in 14 mL tetrahydrofuran and 7 mL ethanol was prepared and placed under a positive nitrogen flow. The mixture was stirred at room temperature for 24 hours. The resulting dark solution was diluted with water (30 mL) and extracted with methylene chloride (30 mL). The organic layer was washed three times with H<sub>2</sub>O, dried over MgSO<sub>4</sub>, and filtered through Celite. The solvent was then dried *in vacuo*. The residue was purified by flash chromatography (eluent: DCM;  $R_f = 0.4$ ) through a silica column, using UV light for TLC visualization. The resulting solution was dried under vacuum, providing 4-pentafluorophenyl(hydroxymethyl) benzene as a white solid (0.705 g, 78%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  7.49 (d, 2H, <u>Ar</u>), 7.42 (d, 2H, <u>Ar</u>), 4.76 (d, 2H, C<u>H</u><sub>2</sub>OH), 2.05 (t, 1H, CH<sub>2</sub>O<u>H</u>). <sup>13</sup>C{<sup>1</sup>H} NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  144.3, 142.3, 140.6,

138.1, 130.5, 127.2, 126.3, 115.8, 64.9. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>): δ -143.3 (q, 2F, -*ortho*), -155.5 (t, 1F, -*para*), -162.2 (m, 2F, -*meta*).

# Synthesis of 4-pentafluorophenyl(bromomethyl) benzene



A flask containing 4-pentafluorophenyl(hydroxymethyl) benzene (1.00 g, 3.65 mmol) was purged with nitrogen and 30 mL of dry methylene chloride was charged into the flask. The solution was placed in ice bath and PBr<sub>3</sub> (346  $\mu$ L, 3.65 mmol) was added *via* syringe. The reaction mixture was stirred overnight, during which time the mixture turned yellow. The resulting mixture was then diluted with 100 mL distilled H<sub>2</sub>O. The organic layer was separated and washed 3 times with saturated NaCl solution. Organic layer was collected and dried over MgSO<sub>4</sub>, then filtered through Celite. Solvent was evaporated and the residue was purified by flash chromatography (hexane/CH<sub>2</sub>Cl<sub>2</sub>, 2:1; R<sub>f</sub> = 0.75) through a silica column, using UV light for TLC visualization. The resulting solution was dried under vacuum, providing 4-pentafluorophenyl(bromomethyl) benzene as a white solid (0.773 g, 63%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  7.53 (d, 2H, <u>Ar</u>), 7.42 (d, 2H, <u>Ar</u>), 4.54 (s, 2H, C<u>H</u><sub>2</sub>Br). <sup>13</sup>C{<sup>1</sup>H} NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  144.3, 140.7, 139.1, 138.0, 130.7, 129.5, 126.6, 115.4, 32.6. <sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>):  $\delta$  -143.1 (q, 2F, *-ortho*), -155.1 (t, 1F, *-para*), -162.0 (m, 2F, *-meta*).

## Synthesis of 3



Compound 1 (75.0 mg, 0.092 mmol) was added to a 10 mL glass microwave vial and transferred out of a nitrogen-filled glovebox, opened to the air, and dissolved in 1.5 mL acetonitrile. N,Ndiisopropylethylamine (0.3 mL, 1.73 mmol) and 4-pentafluorophenyl(bromomethyl) benzene (0.8334 g, 2.47 mmol) were added along with a flea micro stir bar, the vial was sealed with a PTFE/silicone cap, and the mixture was heated at 140 °C with stirring in the microwave for 30 minutes. The volatiles were removed *via* rotary evaporation, and the remaining reagent was eluted first through a slurry-packed silica gel column with 80/20 hexanes/CH<sub>2</sub>Cl<sub>2</sub>, and the pink/purple product mixture was eluted with acetone followed by CH<sub>2</sub>Cl<sub>2</sub>. Note: The eluted fraction containing the reagent ligand can be purified by eluting through a silica column with 90/10 hexanes/CH<sub>2</sub>Cl<sub>2</sub>, and after drying thoroughly it can be used for subsequent synthesis of 3. Recycling the ligand in this manner can minimize unnecessary repetition of ligand synthesis. The volatiles were removed via rotary evaporation, and the remaining charged 2-/1- product mixture was dissolved in 5 mL 90/10 EtOH/MeCN. FeCl<sub>3</sub>·6H<sub>2</sub>O (0.3 g, 1.11 mmol) was added and the mixture was left to stir for 24 hours. Following oxidation, the solvent mixture was removed via rotary evaporation, and a redorange band containing the neutral product was separated from the FeCl<sub>3</sub>·6H<sub>2</sub>O through a slurrypacked silica gel column with  $CH_2Cl_2$ . The  $CH_2Cl_2$  was removed via rotary evaporation and the final neutral product 2 was dried under high vacuum to obtain an isolated yield of 266.5 mg (85%). Compound 2 is a dark red-orange solid. <sup>1</sup>H NMR (500 MHz, CD<sub>2</sub>Cl<sub>2</sub>): δ 7.21 - 7.33 (m, 48H, C<sub>6</sub><u>H</u><sub>4</sub>), 5.50 (s, 24H, OC<u>H</u><sub>2</sub>). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>2</sub>Cl<sub>2</sub>): δ 42.4. <sup>19</sup>F NMR (376 MHz, CD<sub>2</sub>Cl<sub>2</sub>): δ -144.2 (q, 24F, -ortho), -156.5 (t, 12F, -para), -163.4 – -163.5 (m, 24F, -meta). HRMS (Q-Exactive Plus): m/z calculated for C<sub>165</sub>H<sub>72</sub>B<sub>12</sub>F<sub>60</sub>O<sub>12</sub> (M<sup>-</sup>), 3407.5289 Da; found, 3407.5278 Da. X-ray quality crystals of **3** were grown from a cooling solution of boiling 1:1 EtOH:MeOH.



# **Supplementary Table 1. Initial Studies and Reaction Optimization**

<sup>a</sup>Yield determined by <sup>19</sup>F NMR. Tris, tris(hydroxymethyl)aminomethane.

# Synthesis of 2a



**2** (5.0 mg, 0.0020 mmol) and K<sub>2</sub>CO<sub>3</sub> (8.4 mg, 0.061 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150  $\mu$ L anhydrous DMF was added, followed by 1-hexanethiol (3.76  $\mu$ L, 0.027 mmol). The vial was sealed again and set to stir at 400 rpm for 22 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 5 <sup>3</sup>/<sub>4</sub>" glass Pasteur

pipet column was prepared using glass wool and 4" of silica gel, and the pipet was flushed with triethylamine (2X column volumes). The crude product mixture containing **2a** was loaded onto the column with 80/20 hexanes/ethyl acetate (sonication was used to aid dissolution), and the remaining reagent was eluted with 80/20 hexanes/ethyl acetate. A very slightly yellow band containing **2a** was eluted with MeCN, and the fractions containing **2a** (as assessed by TLC) were combined and volatiles were removed *via* rotary evaporation followed by lyophilization overnight to obtain an isolated yield of 5.4 mg (70%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN):  $\delta$  5.42 (br s, 24H, OCH<sub>2</sub>), 3.12 (q, 12H, [(CH<sub>3</sub>CH<sub>2</sub>)<sub>3</sub>NH]<sup>+</sup>), 2.89 – 2.82 (m, 24H, SCH<sub>2</sub>), 1.49 - 1.39 (m, 24H, SCH<sub>2</sub>CH<sub>2</sub>), 1.36 – 1.26 (br m, 24H, S(CH<sub>2</sub>)<sub>2</sub>(CH<sub>2</sub>)<sub>3</sub>CH<sub>3</sub>), 1.24 (t, 18H, [(CH<sub>3</sub>CH<sub>2</sub>)<sub>3</sub>NH]<sup>+</sup>), 1.21 – 1.10 (br m, 48H, S(CH<sub>2</sub>)<sub>2</sub>(CH<sub>2</sub>)<sub>3</sub>CH<sub>3</sub>), 0.83 – 0.74 (m, 36H, S(CH<sub>2</sub>)<sub>5</sub>CH<sub>3</sub>). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>3</sub>CN):  $\delta$  -15.8. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>CN):  $\delta$  -137.4 (br m, 24F, *-meta*<sup>7</sup>), -145.1 (br m, 24F, *-ortho*<sup>7</sup>). MS (LCT Premier): *m/z* calculated for C<sub>156</sub>H<sub>180</sub>B<sub>12</sub>F<sub>48</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2-</sup>), 1836.52 Da; found, 1836.29 Da.

#### Synthesis of 2b



2 (5.0 mg, 0.0020 mmol) and  $K_3PO_4$  (9 mg, 0.042 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150 µL anhydrous DMF was added, followed by thiophenol (2.66 µL, 0.026 mmol). The vial was sealed again and set to stir at 400 rpm for 25 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 5 <sup>3</sup>/<sub>4</sub>" glass Pasteur pipet column was prepared using glass wool and 4" of silica gel, and the pipet was flushed with triethylamine (2X column volumes). The crude product mixture containing **2b** was loaded onto the column with 35/65 ethyl acetate/hexanes (sonication was used to aid dissolution), and the remaining reagent was eluted with 35/65 ethyl acetate/hexanes. A very slightly yellow band containing **2b** was eluted with MeCN, and the fractions containing **2b** (as assessed by TLC) were combined and volatiles were removed *via* rotary evaporation followed by lyophilization overnight to obtain an isolated yield of 6.8 mg (90%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN):  $\delta$  7.22 – 7.14 (br m, 60H, S-<u>Ar</u>), 5.49 (br s, 24H, OC<u>H</u><sub>2</sub>), 3.11 (q, 12H, [(CH<sub>3</sub>CH<sub>2</sub>)<sub>3</sub>NH]<sup>+</sup>), 1.23 (t, 18H, [(C<u>H</u><sub>3</sub>CH<sub>2</sub>)<sub>3</sub>NH]<sup>+</sup>). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>3</sub>CN):  $\delta$  -15.7. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>CN):  $\delta$  -136.4 (m, 24F, -*meta*), -144.1 (m, 24F, -*ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>156</sub>H<sub>84</sub>B<sub>12</sub>F<sub>48</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2-</sup>), 1788.1481 Da; found, 1788.1514 Da.

#### Synthesis of 2c



**2** (5.0 mg, 0.0020 mmol) and K<sub>3</sub>PO<sub>4</sub> (8.1 mg, 0.038 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150  $\mu$ L anhydrous DMF was added, followed by benzyl mercaptan (3.53  $\mu$ L, 0.030 mmol). The vial was sealed again and set to stir at

400 rpm for 24 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for in situ <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 5<sup>3</sup>/<sub>4</sub>" glass Pasteur pipet column was prepared using glass wool and 4" of silica gel, and the pipet was flushed with triethylamine (2X column volumes). The crude product mixture containing 2c was loaded onto the column with 35/65 ethyl acetate/hexanes (sonication was used to aid dissolution), and the remaining reagent was eluted with 35/65 ethyl acetate/hexanes. A very slightly yellow band containing 2c was eluted with MeCN, and the fractions containing 2c (as assessed by TLC) were combined and volatiles were removed via rotary evaporation followed by lyophilization overnight to obtain an isolated yield of 7.4 mg (93.5%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN):  $\delta$  7.20 – 7.04 (br m, 60H, SCH<sub>2</sub>-Ar), 5.39 (br s, 24H, OCH<sub>2</sub>), 4.05 (m, 24H, SCH<sub>2</sub>), 3.11 (q, 12H,  $[(CH_3CH_2)_3NH]^+$ ), 1.23 (t, 18H,  $[(CH_3CH_2)_3NH]^+$ ). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>3</sub>CN): δ-15.8. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>CN): δ-136.8 (m, 24F, -meta), -144.8 (m, 24F, -ortho). HRMS (Q-Exactive Plus): m/z calculated for  $C_{168}H_{108}B_{12}F_{48}O_{12}S_{12}$  (M<sup>2-</sup>), 1872.2420 Da; found, 1872.2469 Da.

# Synthesis of 2d



2 (5.0 mg, 0.0020 mmol) and  $K_3PO_4$  (10.4 mg, 0.049 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred

into the glovebox. In the glovebox, the vial was opened and 150  $\mu$ L anhydrous DMF was added, followed by 2-mercaptoethanol (2.26 µL, 0.032 mmol). The vial was sealed again and set to stir at 400 rpm for 24 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for in situ <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ*<sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the crude product mixture containing 2d was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated vield of 2.6 mg (40 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 5.50 (br s, 24H, OCH<sub>2</sub>), 3.64 (t, 24H, CH<sub>2</sub>CH<sub>2</sub>OH), 3.00 (t, SCH<sub>2</sub>CH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD): δ -15.7. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD): δ -137.6 - -137.7 (m, 24F, -meta), -145.1 – -145.2 (m, 24F, -ortho). HRMS (Q-Exactive Plus): m/z calculated for C<sub>108</sub>H<sub>84</sub>B<sub>12</sub>F<sub>48</sub>O<sub>24</sub>S<sub>12</sub> (M<sup>2-</sup>), 1596.1176 Da; found, 1596.1233 Da.

Synthesis of 2e



2 (5.0 mg, 0.0020 mmol) and  $K_3PO_4$  (10.2 mg, 0.048 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150 µL anhydrous DMF was added,

followed by thioglycerol (3.12  $\mu$ L, 0.036 mmol). The vial was sealed again and set to stir at 400 rpm for 24 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for in situ <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and in situ <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the crude product mixture containing 2e was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried via rotary evaporation, and subjected to characterization via <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried via rotary evaporation to obtain an isolated yield of 2.2 mg (30 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 5.50 (br s, 24H, O<u>CH</u><sub>2</sub>), 3.69 – 3.64 (m, 12H, SCH<sub>2</sub>C<u>H</u>(OH)), 3.60 – 3.53 (m, 24H, CH(OH)CH<sub>2</sub>OH), 3.07 – 2.93 (m, 24H, SCH<sub>2</sub>CH(OH)). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD): δ -15.6. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD): δ -137.5 – -137.6 (m, 24F, -meta), -145.1 (m, 24F, -ortho). HRMS (Q-Exactive Plus): m/z calculated for C<sub>120</sub>H<sub>108</sub>B<sub>12</sub>F<sub>48</sub>O<sub>36</sub>S<sub>12</sub> (M<sup>2-</sup>), 1776.1810 Da; found, 1776.1894 Da.

#### Synthesis of 2f



2 (5.0 mg, 0.0020 mmol) and  $K_2CO_3$  (2.6 mg, 0.019 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150 µL anhydrous DMF was added,

followed by cysteamine (3.7 mg, 0.048 mmol). The vial was sealed again and set to stir at 400 rpm for 23 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in 40/60 MeOH/MeCN (23 cm packed height), and the crude product mixture containing **2f** was loaded onto the column with 40/60 MeOH/MeCN. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 3.2 mg (49 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  5.51 (br s, 24H, OCH<sub>2</sub>), 2.94 (t, 24H, SCH<sub>2</sub>CH<sub>2</sub>), 2.72 (t, CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.4. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -137.6 (m, 24F, *-meta*), -144.4 – -144.6 (m, 24F, *-ortho*). MS (LCT Premier): *m/z* calculated for C<sub>108</sub>H<sub>96</sub>B<sub>12</sub>F<sub>48</sub>N<sub>12</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2</sup>), 1590.21 Da; found, 1590.07 Da.

#### Synthesis of 2i



**2** (8 mg, 0.0032 mmol) and K<sub>3</sub>PO<sub>4</sub> (16.6 mg, 0.078 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 240  $\mu$ L anhydrous DMF was added, followed by mPEGthiol<sub>356</sub> (20.63  $\mu$ L, 0.064 mmol). The vial was sealed again and set to stir at 400 rpm for

28 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the crude product mixture containing **2i** was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 16.9 mg (81 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  5.51 (br s, 24H, OCH<sub>2</sub>), 3.63 – 3.50 (m, 312H, SCH<sub>2</sub>CH<sub>2</sub>O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>6</sub>), 3.35 – 3.33 (m, 36H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>6</sub>CH<sub>3</sub>), 3.08 (t, 24H, SCH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.7. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -137.2 – -137.3 (m, 24F, *-meta*), -144.8 (m, 24F, *-ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>264</sub>H<sub>396</sub>B<sub>12</sub>F<sub>48</sub>O<sub>96</sub>S<sub>12</sub> (M<sup>2-</sup>), 3265.1552 Da; found, 3265.1444 Da.

#### Synthesis of j



In a round bottom flask, mPEG<sub>750</sub> (7.50 g, 10.00 mmol) and CBr<sub>4</sub> (3.98 g, 12.00 mmol) were dissolved in 40 mL of acetonitrile. To the stirring solution, PPh<sub>3</sub> (3.15 g, 6.00 mmol) was added in small portions over 30 minutes. The mixture was then left stirring at room temperature for 4 hours. After 4 hours, the solvent was then removed *in vacuo* and the resulting yellow-orange oil was dissolved in 20 mL of H<sub>2</sub>O and left at 4 °C overnight, producing a white precipitate. The mixture was filtered through Celite\* on a glass frit and the filtrate was washed twice with 5 mL of

toluene. The aqueous layer was dried *in vacuo* to yield the desired product as an orange oil (7.08 g, 87%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  3.55 – 3.51 (m, 62H, CH<sub>2</sub>O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>), 3.43 (m, 2H, BrCH<sub>2</sub>), 3.26 (s, 3H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>CH<sub>3</sub>).

\*Celite was pretreated on the frit by washing with 30 mL of H<sub>2</sub>O before the mixture was filtered.

#### 2. Synthesis of j-SAc



To a solution of **j-Br** (1.07 g, 1.32 mmol) in 35 mL of ethanol, potassium thioacetate (0.20 g, 1.75 mmol) was added in one portion. The mixture was refluxed at 120 °C for 5 hours. The resulting suspension was filtered through Celite and the filtrate was dried under vacuum, affording a brown oil. The oil was dissolved in 40 mL of chloroform and the organic phase was washed twice with H<sub>2</sub>O. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and filtered through Celite. The solvent was removed *in vacuo*, providing **j-SAc** (0.64 g, 74%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  3.64 – 3.61 (m, 62H, CH<sub>2</sub>O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>), 3.36 (s, 3H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>CH<sub>3</sub>), 3.07 (t, 2H, SCH<sub>2</sub>), 2.32 (s, 3H, SCOCH<sub>3</sub>).

3. Synthesis of j



**j-SAc** (405 mg, 0.5 mmol) was charged with 5 mL of 1M HCl and was refluxed at 110 °C for 2 hours under a blanket of Ar. The solvent was removed *in vacuo*. The residue was dissolved in 10 mL of DCM and the organic phase was washed twice with water. The organic layer was separated and dried over Na<sub>2</sub>SO<sub>4</sub> and filtered through Celite. The solution was dried under vacuum to yield the desired product as a brown oil (319 mg, 83%). Product was stored under inert atmosphere. <sup>1</sup>H

NMR (400 MHz,  $CD_2Cl_2$ ):  $\delta$  3.58 – 3.59 (m, 62H,  $CH_2O(CH_2CH_2O)_{15}$ ), 3.32 (s, 3H,  $(CH_2CH_2O)_{15}CH_3$ ), 2.67 (dt, 2H,  $SHCH_2$ ), 1.61 (t, 1H,  $SHCH_2$ ).

Synthesis of 2j



2 (5.0 mg, 0.0020 mmol) and K<sub>3</sub>PO<sub>4</sub> (19.2 mg, 0.090 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N2 three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150 µL anhydrous DMF was added, followed by mPEGthiol<sub>766</sub> (48.1 µL, 0.069 mmol). The vial was sealed again and set to stir at 400 rpm for 24 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and in situ<sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex G50 medium in water (23 cm packed height), and the crude product mixture containing 2j was loaded onto the column with water. 15 1-2 mL fractions were collected, dried *via* lyophilization, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* lyophilization to obtain an isolated yield of 4.4 mg (19%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 5.50 (br s, 24H, OCH<sub>2</sub>), 3.63 – 3.53 (m, 744H, SCH<sub>2</sub>CH<sub>2</sub>O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>), 3.35 (m, 36H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>CH<sub>3</sub>), 3.08 (t, 24H, SCH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD): δ -16.0. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD): δ -137.2 (m, 24F, -meta), -144.8 (m, 24F, -ortho).

Synthesis of 2k



2 (5.0 mg, 0.0020 mmol) and K<sub>3</sub>PO<sub>4</sub> (13.4 mg, 0.063 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N2 three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150 µL anhydrous DMF was added, followed by mPEGthiol<sub>2000</sub> (101.0 mg, 0.051 mmol). The vial was sealed again and set to stir at 400 rpm for 24 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for in situ <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and in situ <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex G50 medium in water (23 cm packed height), and the crude product mixture containing 2k was loaded onto the column with water. 15 1-2 mL fractions were collected, dried via lyophilization, and subjected to characterization via <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* lyophilization to obtain an isolated yield of 21.5 mg (41 %). <sup>1</sup>H NMR (400 MHz. CD<sub>3</sub>OD):  $\delta$  5.50 (br s, 24H, OCH<sub>2</sub>), 3.82 – 3.45 (m, 2100H, SCH<sub>2</sub>CH<sub>2</sub>(CONH)CH<sub>2</sub>CH<sub>2</sub>O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>42</sub>), 3.36 (s, 36H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>42</sub>CH<sub>3</sub>), 3.09 (t, 24H, SCH<sub>2</sub>CH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD): δ -16.0. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD): δ -137.0 - -137.1 (m, 24F, *-meta*), -144.8 (m, 24F, *-ortho*). GPC trace of **2k** measured in water with 0.02% NaN<sub>3</sub> at 12 °C gives a Đ (polydispersity index) of 1.003 (see Fig. 3c in main text).

Synthesis of 21



2 (5.0 mg, 0.0020 mmol) and K<sub>3</sub>PO<sub>4</sub> (13.0 mg, 0.061 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N2 three times before being transferred into the glovebox. In the glovebox, the vial was opened and 150 µL anhydrous DMF was added, followed by 1-thio- $\beta$ -D-glucose tetraacetate (25.0 mg, 0.069 mmol). The vial was sealed again and set to stir at 400 rpm for 24 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and in situ <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. The resulting residue was treated with NaOMe (6.0 mg, 0.11 mmol) in 1 mL MeOH for 2 hours. The volatiles were removed *via* rotary evaporation. A 1.25 cm x 35 cm glass column was packed with Sephadex G50 medium in water (23 cm packed height), and the crude product mixture containing 21 was loaded onto the column with water. 15 1-2 mL fractions were collected, dried via lyophilization, and subjected to characterization via <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried via lyophilization to obtain an isolated yield of 1.6 mg (17 %). <sup>1</sup>H NMR (400 MHz,  $D_2O$ ):  $\delta$  5.64 – 5.45 (br s, 24H, OCH<sub>2</sub>), 4.03 – 3.20 (m, 84H, SCHCH<sub>2</sub>OH(CHOH)<sub>3</sub>CHO). <sup>11</sup>B NMR (128 MHz, D<sub>2</sub>O): δ -16.3. <sup>19</sup>F NMR (376 MHz, D<sub>2</sub>O): δ -134.3 - -135.6 (m, 24F, -meta), -143.5 (m, 24F, -

*ortho*). HRMS (Q-Exactive Plus): m/z calculated for C<sub>156</sub>H<sub>156</sub>B<sub>12</sub>F<sub>48</sub>O<sub>72</sub>S<sub>12</sub> (M<sup>2-</sup>), 2304.2772 Da; found, 2304.2769 Da.

Synthesis of 3a



3 (10.0 mg, 0.0029 mmol) and K<sub>2</sub>CO<sub>3</sub> (22.0 mg, 0.159 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N2 three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300 µL anhydrous DMF was added, followed by 1-hexanethiol (5.42 µL, 0.038 mmol). The vial was sealed again and set to stir at 400 rpm for 7 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for in situ <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and in situ <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 5 <sup>3</sup>/<sub>4</sub>" glass Pasteur pipet column was prepared using glass wool and 4" of silica gel, and the pipet was flushed with triethylamine (2X column volumes). The crude product mixture containing **3a** was loaded onto the column with 80/20 hexanes/ethyl acetate (sonication was used to aid dissolution), and the remaining reagent was eluted with 80/20 hexanes/ethyl acetate. A very slightly yellow band containing 3a was eluted with MeCN, and the fractions containing 3a (as assessed by TLC) were combined and volatiles were removed via rotary evaporation followed by lyophilization overnight to obtain an isolated yield of 12.2 mg (87%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN):  $\delta$  7.64 – 7.50 (br m, 24H, OCH<sub>2</sub>-Ar), 7.25 - 7.15 (br m, 24H, OCH<sub>2</sub>-Ar), 5.60 (br s, 24H, OCH<sub>2</sub>), 3.06 (q, 12H,

[(CH<sub>3</sub>C<u>H<sub>2</sub></u>)<sub>3</sub>NH]<sup>+</sup>), 2.93 (t, 24H, SC<u>H<sub>2</sub></u>), 1.61 - 1.49 (m, 24H, SCH<sub>2</sub>C<u>H<sub>2</sub></u>), 1.44 – 1.34 (br m, 24H, S(CH<sub>2</sub>)<sub>2</sub>(C<u>H<sub>2</sub></u>)<sub>3</sub>CH<sub>3</sub>), 1.30 – 1.21 (br m, 48H, S(CH<sub>2</sub>)<sub>2</sub>(C<u>H<sub>2</sub></u>)<sub>3</sub>CH<sub>3</sub>), 1.18 (t, 18H, [(C<u>H<sub>3</sub>CH<sub>2</sub></u>)<sub>3</sub>NH]<sup>+</sup>), 0.89 – 0.80 (m, 36H, S(CH<sub>2</sub>)<sub>5</sub>C<u>H<sub>3</sub></u>). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>3</sub>CN): δ -15.1. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>CN): δ -136.7 (q, 24F, *-meta*), -145.2 (q, 24F, *-ortho*). HRMS (Q-Exactive Plus): m/z calculated for C<sub>228</sub>H<sub>228</sub>B<sub>12</sub>F<sub>48</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2-</sup>), 2292.7115 Da; found, 2292.7157 Da.

#### Synthesis of 3b



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (18.9 mg, 0.089 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by thiophenol (3.89  $\mu$ L, 0.038 mmol). The vial was sealed again and set to stir at 400 rpm for 7 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 5 <sup>3</sup>/<sub>4</sub>" glass Pasteur pipet column was prepared using glass wool and 4" of silica gel, and the pipet was flushed with triethylamine (2X column volumes). The crude product mixture containing **3b** was loaded onto the column with 35/65 ethyl acetate/hexanes (sonication was used to aid dissolution), and the remaining reagent was eluted with MeCN, and the fractions containing **3b** (as assessed by TLC) were

combined and volatiles were removed *via* rotary evaporation followed by lyophilization overnight to obtain an isolated yield of 11.7 mg (85%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN):  $\delta$  7.65 – 7.48 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.34 – 7.20 (br m, 60H and 24H, S-<u>Ar</u> and OCH<sub>2</sub>-<u>Ar</u>), 5.61 (br s, 24H, OC<u>H<sub>2</sub></u>), 3.09 (q, 12H, [(CH<sub>3</sub>C<u>H<sub>2</sub></u>)<sub>3</sub>NH]<sup>+</sup>), 1.21 (t, 18H, [(C<u>H<sub>3</sub>CH<sub>2</sub></u>)<sub>3</sub>NH]<sup>+</sup>). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>3</sub>CN):  $\delta$  -15.1. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>CN):  $\delta$  -135.9 (m, 24F, -*meta*), -145.2 (m, 24F, -*ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>228</sub>H<sub>132</sub>B<sub>12</sub>F<sub>48</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2-</sup>), 2244.3359 Da; found, 2244.3381 Da.

# Synthesis of 3c



**3** (10.0 mg, 0.0029 mmol) and  $K_3PO_4$  (22.5 mg, 0.106 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300 µL anhydrous DMF was added, followed by benzyl mercaptan (4.48 µL, 0.038 mmol). The vial was sealed again and set to stir at 400 rpm for 5 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 5 <sup>3</sup>/<sub>4</sub>" glass Pasteur pipet column was prepared using glass wool and 4" of silica gel, and the pipet was flushed with triethylamine (2X column volumes). The crude product mixture containing **3c** was loaded onto the column with 35/65 ethyl acetate/hexanes (sonication was used to aid dissolution), and the

remaining reagent was eluted with 35/65 ethyl acetate/hexanes. A very slightly yellow band containing **3c** was eluted with MeCN, and the fractions containing **3c** (as assessed by TLC) were combined and volatiles were removed *via* rotary evaporation followed by lyophilization overnight to obtain an isolated yield of 11.6 mg (81%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>CN):  $\delta$  7.59 – 7.52 (br d, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.26 – 7.15 (br m, 60H and 24H, SCH<sub>2</sub>-<u>Ar</u> and OCH<sub>2</sub>-<u>Ar</u>), 5.60 (br s, 24H, OC<u>H<sub>2</sub>), 4.11 (br s, 24H, SC<u>H<sub>2</sub></u>), 3.06 (q, 12H, [(CH<sub>3</sub>C<u>H<sub>2</sub>)<sub>3</sub>NH]<sup>+</sup>), 1.18 (t, 18H, [(CH<sub>3</sub>CH<sub>2</sub>)<sub>3</sub>NH]<sup>+</sup>). <sup>11</sup>B{<sup>1</sup>H} NMR (128 MHz, CD<sub>3</sub>CN):  $\delta$  -15.1. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>CN):  $\delta$  -135.9 (q, 24F, *-meta*), -145.1 (q, 24F, *-ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>240</sub>H<sub>156</sub>B<sub>12</sub>F<sub>48</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2</sup>-), 2328.9298 Da; found, 2328.9363 Da.</u></u>

#### Synthesis of 3d



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (12.3 mg, 0.058 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by 2-mercaptoethanol (2.69  $\mu$ L, 0.038 mmol). The vial was sealed again and set to stir at 400 rpm for 4 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed)

height), and the crude product mixture containing **3d** was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 10.0 mg (81 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  7.61 – 7.45 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.24 – 7.13 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.65 (br m, 24H, OC<u>H<sub>2</sub></u>), 3.73 (t, 24H, CH<sub>2</sub>C<u>H<sub>2</sub></u>OH), 3.10 (t, SC<u>H<sub>2</sub></u>CH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.1. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -136.8 – -136.9 (m, 24F, *-meta*), -145.4 – -145.5 (m, 24F, *-ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>180</sub>H<sub>132</sub>B<sub>12</sub>F<sub>48</sub>O<sub>24</sub>S<sub>12</sub> (M<sup>2</sup>), 2052.3054 Da; found, 2052.3080 Da.

#### Synthesis of 3e



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (13.1 mg, 0.062 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by thioglycerol (3.30  $\mu$ L, 0.038 mmol). The vial was sealed again and set to stir at 400 rpm for 4 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the

crude product mixture containing **3e** was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 7.9 mg (59 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  7.59 – 7.45 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.23 – 7.16 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.64 – 5.60 (br m, 24H, OCH<sub>2</sub>), 3.78 – 3.72 (m, 12H, SCH<sub>2</sub>C<u>H</u>(OH)), 3.65 – 3.57 (m, 24H, CH(OH)C<u>H<sub>2</sub>OH), 3.17 – 3.02 (m, 24H, SCH<sub>2</sub>CH(OH)). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.1. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -136.6 – -136.7 (m, 24F, *-meta*), -145.5 (m, 24F, *-ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>192</sub>H<sub>156</sub>B<sub>12</sub>F<sub>48</sub>O<sub>36</sub>S<sub>12</sub> (M<sup>2-</sup>), 2232.3688 Da; found, 2232.3752 Da.</u>

#### Synthesis of 3f



**3** (10.0 mg, 0.0029 mmol) and K<sub>2</sub>CO<sub>3</sub> (8.1 mg, 0.059 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by cysteamine (4.9 mg, 0.064 mmol). The vial was sealed again and set to stir at 400 rpm for 4 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in 40/60 MeOH/MeCN (23 cm packed height),

and the crude product mixture containing **3f** was loaded onto the column with 40/60 MeOH/MeCN. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 4.0 mg (33 %). <sup>1</sup>H NMR (400 MHz, 33/67 CD<sub>3</sub>OD/CD<sub>3</sub>CN):  $\delta$  7.55 – 7.52 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.21 – 7.18 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.60 – 5.54 (br m, 24H, OCH<sub>2</sub>), 2.95 (t, 24H, SCH<sub>2</sub>CH<sub>2</sub>), 2.70 (t, CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, 33/67 CD<sub>3</sub>OD/CD<sub>3</sub>CN):  $\delta$  -15.2. <sup>19</sup>F NMR (376 MHz, 33/67 CD<sub>3</sub>OD/CD<sub>3</sub>CN):  $\delta$  -136.0 – -136.5 (m, 24F, -meta), -145.1 – -145.6 (m, 24F, -ortho). MS (LCT Premier): *m/z* calculated for C<sub>180</sub>H<sub>144</sub>B<sub>12</sub>F<sub>48</sub>N<sub>12</sub>O<sub>12</sub>S<sub>12</sub> (M<sup>2</sup>-), 2046.40 Da; found, 2046.31 Da.

# Synthesis of 3g



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (18.7 mg, 0.088 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by *N*-(*tert*-Butoxycarbonyl)-L-cysteine methyl ester (8.16  $\mu$ L, 0.040 mmol). The vial was sealed again and set to stir at 400 rpm for 3 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for

solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the crude product mixture containing **3g** was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 8.8 mg (49 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  7.52 (d, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.19 (d, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.63 (br s, 24H, OCH<sub>2</sub>), 4.37 – 4.34 (br m, 12H, SCH<sub>2</sub>C<u>H</u>), 3.69 (m, 36H, OC<u>H<sub>3</sub></u>), 3.49 – 3.44 (br m, 24H, SC<u>H<sub>2</sub></u>), 1.35 – 1.33 (m, 108H, C(C<u>H<sub>3</sub></u>)<sub>3</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.1. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -135.9 – -136.0 (m, 24F, *-meta*), -144.8 – 145.1 (m, 24F, *-ortho*). HRMS (Q-Exactive Plus): *m/z* calculated for C<sub>264</sub>H<sub>264</sub>B<sub>12</sub>F<sub>48</sub>N<sub>12</sub>O<sub>60</sub>S<sub>12</sub> (M<sup>2-</sup>), 2994.7487 Da; found, 2994.7404 Da.

### Synthesis of 3h



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (56.1 mg, 0.264 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by unprotected C-A-G·TFA (synthesized using conventional Fmoc solid-phase peptide synthesis protocol<sup>8</sup>) (17.8 mg, 0.049 mmol) and isopropoxyltrimethylsilane (18.8  $\mu$ L, 0.106 mmol). The vial was sealed again and set to stir at 400 rpm for 6 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR

spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in H<sub>2</sub>O/ACN (23 cm packed height), and the crude product mixture containing **3h** was loaded onto the column with H<sub>2</sub>O/ACN. 15 1-2 mL fractions were collected, dried *via* lyophilization, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* lyophilization to obtain an isolated yield of 5.3 mg (29 %). <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O/CD<sub>3</sub>CN):  $\delta$  7.44 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.09 – 7.08 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.50 (br s, 24H, O-C<u>H<sub>2</sub>), 3.77 – 3.68 (br m, 24H, (CONH)C<u>H<sub>2</sub>(CONH<sub>2</sub>))</u>, 3.48 – 3.45 (br t, 12H, SCH<sub>2</sub>C<u>H</u>), 3.15 – 3.10 (br m, 24H, SC<u>H<sub>2</sub>), 1.26 = 1.24 (d, 36H, CC<u>H<sub>3</sub></u>). <sup>11</sup>B NMR (128 MHz, D<sub>2</sub>O/CD<sub>3</sub>CN):  $\delta$  -15.8. <sup>19</sup>F NMR (376 MHz, D<sub>2</sub>O/CD<sub>3</sub>CN):  $\delta$  -135.4 – -135.5 (m, 24F, *-meta*), -144.7 – -144.8 (m, 24F, *-ortho*). MS (LCT Premier): *m/z* calculated for C<sub>252</sub>H<sub>252</sub>B<sub>12</sub>F<sub>48</sub>N<sub>48</sub>O<sub>48</sub>S<sub>12</sub> (M<sup>2-</sup>), 3072.79 Da; found, 3072.60 Da.</u></u>

#### Synthesis of 3i



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (8.5 mg, 0.040 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by mPEGthiol<sub>356</sub> (12.27  $\mu$ L, 0.038 mmol). The vial was sealed again and set to stir at 400 rpm for 5 hours. The vial was transferred out of the glovebox, and its contents were transferred

into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the crude product mixture containing **3i** was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 17.1 mg (78 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  7.64 – 7.46 (br m, 24H, OCH<sub>2</sub>-A<u>r</u>), 7.26 – 7.18 (br m, 24H, OCH<sub>2</sub>-A<u>r</u>), 5.65 – 5.61 (br m, 24H, OCH<sub>2</sub>), 3.70 (t, 24H, SCH<sub>2</sub>C<u>H<sub>2</sub></u>), 3.62 – 3.44 (m, 288H, SCH<sub>2</sub>CH<sub>2</sub>O(C<u>H<sub>2</sub>CH<sub>2</sub>O)<sub>6</sub>), 3.30 – 3.28 (m, 36H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>6</sub>C<u>H<sub>3</sub></u>), 3.14 (t, 24H, SC<u>H<sub>2</sub></u>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.3. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -136.4 – -136.5 (m, 24F, *-meta*), -145.3 (m, 24F, *-ortho*). HRMS (Q-Exactive Plus): *m*/z calculated for C<sub>336</sub>H<sub>444</sub>B<sub>12</sub>F<sub>48</sub>O<sub>96</sub>S<sub>12</sub> (M<sup>2-</sup>), 3721.3430 Da; found, 3721.3395 Da.</u>

#### Synthesis of 3j



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (32.0 mg, 0.151 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by mPEGthiol<sub>766</sub> (44.1  $\mu$ L, 0.063 mmol). The vial was sealed again and set to stir at 400

rpm for 4 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex LH20 medium in MeOH (23 cm packed height), and the crude product mixture containing **3j** was loaded onto the column with MeOH. 15 1-2 mL fractions were collected, dried *via* rotary evaporation, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* rotary evaporation to obtain an isolated yield of 7.7 mg (21 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 7.57 – 7.55 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.18 – 7.16 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.67 – 5.62 (br m, 24H, OC<u>H<sub>2</sub>), 3.72 (t, 24H, SCH<sub>2</sub>C<u>H<sub>2</sub>), 3.64 – 3.51 (m, 744H, SCH<sub>2</sub>CH<sub>2</sub>O(C<u>H<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>), 3.33</u> (m, 36H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>15</sub>C<u>H<sub>3</sub>), 3.19 – 3.16 (t, 24H, SCH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD): δ -15.3. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD): δ -136.1 – -136.4 (m, 24F, *-meta*), -145.1 (m, 24F, *-ortho*).</u></u></u>

#### Synthesis of 3k



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (27.0 mg, 0.127 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by mPEGthiol<sub>2000</sub> (85.0 mg, 0.043 mmol). The vial was sealed again and set to stir at 400 rpm for 20 hours. The vial was transferred out of the glovebox, and its contents were transferred

into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and *in situ* <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. A 1.25 cm x 35 cm glass column was packed with Sephadex G50 medium in water (23 cm packed height), and the crude product mixture containing **3k** was loaded onto the column with water. 15 1-2 mL fractions were collected, dried *via* lyophilization, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. The pure product fractions as indicated by NMR spectroscopy were combined and dried *via* lyophilization to obtain an isolated yield of 43.2 mg (54 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  7.64 – 7.47 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.20 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.66 (br m, 24H, OCH<sub>2</sub>), 3.92 – 3.44 (m, 2100H, SCH<sub>2</sub>CH<sub>2</sub>(CON<u>H</u>)CH<sub>2</sub>CH<sub>2</sub>O(CH<sub>2</sub>CH<sub>2</sub>O)<sub>42</sub>), 3.35 (s, 36H, (CH<sub>2</sub>CH<sub>2</sub>O)<sub>42</sub>CH<sub>3</sub>), 3.19 (t, 24H, SCH<sub>2</sub>-CH<sub>2</sub>). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD):  $\delta$  -15.4. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -136.2 – -137.3 (m, 24F, *-meta*), -145.0 – -145.5 (m, 24F, *-ortho*). GPC trace of **3k** measured in water with 0.02% NaN<sub>3</sub> at 12 °C gives a D (polydispersity index) of 1.081 (see Fig. 3c in main text).

#### Synthesis of 31



**3** (10.0 mg, 0.0029 mmol) and K<sub>3</sub>PO<sub>4</sub> (18.7 mg, 0.088 mmol) were added along with a flea micro stir bar to a 4-mL glass vial, which was then sealed with a PTFE/silicone cap under ambient conditions. The vial was then purged and backfilled with N<sub>2</sub> three times before being transferred into the glovebox. In the glovebox, the vial was opened and 300  $\mu$ L anhydrous DMF was added, followed by 1-thio- $\beta$ -D-glucose tetraacetate (16.4 mg, 0.045 mmol). The vial was sealed again and
set to stir at 400 rpm for 5 hours. The vial was transferred out of the glovebox, and its contents were transferred into an NMR tube for *in situ* <sup>19</sup>F NMR spectroscopy to ensure nearly quantitative conversion and in situ <sup>11</sup>B NMR spectroscopy to ensure structural integrity of the cluster. The crude mixture was then transferred into a 20-mL glass vial and lyophilized for solvent removal. The resulting residue was treated with NaOMe (3.8 mg, 0.07 mmol) in 1 mL MeOH for 2 hours. The volatiles were removed via rotary evaporation. The crude product mixture containing 31 was dissolved in water and adjusted to pH 7.3 using 3M HCl. This mixture was then centrifuged 5 times – after each of the first 4 centrifugation periods, the supernatant was removed by pipet and more water was added, after the 5<sup>th</sup> centrifugation period, the supernatant was removed and the precipitate was dried *via* lyophilization, and subjected to characterization *via* <sup>1</sup>H, <sup>11</sup>B, and <sup>19</sup>F NMR spectroscopy. This pure product as indicated by NMR spectroscopy was dried via lyophilization to obtain an isolated yield of 5.3 mg (32 %). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD):  $\delta$  7.62 – 7.46 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 7.27 – 7.17 (br m, 24H, OCH<sub>2</sub>-<u>Ar</u>), 5.65 – 5.56 (br m, 24H, OCH<sub>2</sub>), 3.77 – 3.33, 3.28 (m, 84H, SCHCH<sub>2</sub>OH(CHOH)<sub>3</sub>CHO). <sup>11</sup>B NMR (128 MHz, CD<sub>3</sub>OD): δ -15.4. <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD): δ -135.4 - -135.5 (m, 24F, -meta), -145.5 - -145.6 (m, 24F, -ortho). HRMS (Q-Exactive Plus): m/z calculated for C<sub>228</sub>H<sub>204</sub>B<sub>12</sub>F<sub>48</sub>O<sub>72</sub>S<sub>12</sub> (M<sup>3-</sup>), 1840.3100 Da; found, 1840.3178 Da.

















	Current Data Parameters
	NAME $D_{12}(O, D_b DED_m) 12$ also $C_{2m}$
	NAME D12(0-PHPFDH)12 aka Genz
	EXPNO 91
	PROCNO I
	F2 Acquisition Parameters
	Doto 20151220
	Date20131220
	Internet 14.27
	INSTRUM av400
	PROBHD 5 mm PABBO BB/
	PULPROG zg30
	TD 52882
	SOLVENT CD2Cl2
	NS 32
	DS 0
	SWH 8012.820 Hz
	FIDRES 0.151523 Hz
	AO 3.2998369 sec
	RG 155.85
	DW 62 400 usec
	DE 6.50 usec
	TE 299.0 K
	D1 = 20000000  sec
	TD0 1
	100 1
	======= CHANNEL f1 =======
	SFO1 400.1324008 MHz
	NUC1 1H
	P1 15.00 usec
	PLW1 13.0000000 W
	F2 - Processing parameters
	SI 65536
	SF 400.1300203 MHz
	WDW EM
	SSB 0
	LB 0.30 Hz
	GB 0
	PC 1.00
	10 1.00
ppm	



#### <sup>11</sup>B {<sup>1</sup>H} NMR













PhPFB #1 RT: 0.01 AV: 1 NL: 6.03E4 T: FTMS - p ESI Full ms [1500.00-4000.00]





#### in situ <sup>11</sup>B NMR







Small impurities are present due to the commercial 1-hexanethiol used (95% pure).





<sup>11</sup>B {<sup>1</sup>H} NMR



													Current Data Parameters NAME Feb23-2016
													$\begin{array}{llllllllllllllllllllllllllllllllllll$
													======   CHANNEL [1]     SF01   128.3776052 MHz     NUC1   11B     P1   10.00 usec     PLW1   52.0000000 W
													======================================
 												·····	F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB 0
60	50	<b>40</b>	30	20	10	0	-10	-20	-30	-40	-50	ppm	rc 1.40



Small impurities are present due to the commercial 1-hexanethiol used (95% pure).



# Waters Mass Spec

high mz scan 2a 19 (1.060) Cm (9:36)





# Waters Mass Spec

high mz scan 2a 19 (1.060) Cm (9:36)





### *in situ* <sup>11</sup>**B** NMR



			-14.92			Current Data ParametersNAME0129EXPNO221PROCNO1F2 - Acquisition ParametersDate_20160129Time20.43INSTRUMav400PROBHD 5 mm PABBO BB/PULPROGzgTD5096SOLVENTNoneNS1024DS0SWH51020.406 HzFIDRES10.011854 HzAQ0.0499408 secRG189.85DW9.800 usecDE6.50 usecTE299.0 KDI0.05000000 secTDO1=====CHANNEL f1 ======SFO1128.3776052 MHzNUC111BP110.00 usecPLW152.00000000 WF2 - Processing parametersSI32768
 20	 	 		 	 	ST   128.37/0101 MHz     WDW   EM     SSB   0     LB   10.00 Hz     GB   0     PC   1.40





 $\begin{array}{c} & & \\ & &$ 

<sup>11</sup>B {<sup>1</sup>H} NMR



													Current Data Parameters NAME Feb03-2016 EXPNO 100 PROCNO 1
													F2 - Acquisition Parameters   Date_ 20160203   Time 15.06   INSTRUM av400   PROBHD 5 mm PABBO BB/   PULPROG zgdc.js   TD 5096   SOLVENT CD3CN   NS 1024   DS 0   SWH 51020.406 Hz   FIDRES 10.011854 Hz   AQ 0.0499408 sec   RG 189.85   DW 9.800 usec   DE 6.50 usec   TE 299.0 K   D1 0.03000000 sec   D11 0.03000000 sec   TD0 1
													======================================
													Emergence   CHANNEL f2     SFO2   400.1324008 MHz     NUC2   1H     CPDPRG[2   waltz16     PCPD2   90.00 usec     PLW2   13.00000000 W     PLW12   0.36111000 W
 													F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB 0
 60	50	<b>40</b>	30	20	10	0	-10	-20	-30	-40	-50	ppm	PC 1.40









2b 1.25-4k #1-16 RT: 0.01-0.14 AV: 16 NL: 4.68E6 T: FTMS - p ESI Full ms [1250.00-4000.00]





*in situ* <sup>11</sup>**B** NMR



										Current Data Paramatara
										Current Data ParametersNAME0201EXPNO131PROCNO1F2 - Acquisition ParametersDate_20160201Time20.18INSTRUMav400PROBHD 5 mm PABBO BB/PULPROGzgTD5096SOLVENTNoneNS1024DS0SWH51020.406 HzFIDRES10.011854 HzAQ0.0499408 secRG189.85DW9.800 usecDE6.50 usecTE299.0 KD10.05000000 secTD01
 										====== CHANNEL f1 ======   SF01 128.3776052 MHz   NUC1 11B   P1 10.00 usec   PLW1 52.00000000 W   F2 - Processing parameters SI   SI 32768   SF 128.3776161 MHz   WDW EM   SSB 0   LB 10.00 Hz   GB 0   DC 140
 30	20	10	0	-10	-20	-30	-40	-50	ppm	rC 1.40







<sup>11</sup>B {<sup>1</sup>H} NMR



								0/·Cl					Current Data Parameters     NAME   Feb03-2016     EXPNO   90     PROCNO   1     F2 - Acquisition Parameters   Date_     Date_   20160203     Time   14.49     INSTRUM   av400     PROBHD   5 mm PABBO BB/     PULPROG   zgdc.js     TD   5096     SOLVENT   CD3CN     NS   1024     DS   0     SWH   51020.406 Hz     FIDRES   10.011854 Hz     AQ   0.0499408 sec     RG   189.85     DW   9.800 usec     DE   6.50 usec     TE   299.0 K     D1   0.050000000 sec     D10   1     ====================================
													Example   CHANNEL f2     SF02   400.1324008 MHz     NUC2   1H     CPDPRG[2   waltz16     PCPD2   90.00 usec     PLW2   13.00000000 W     PLW12   0.36111000 W
 												~~~~~	F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB 0
 <b>60</b>	50	<b>40</b>	30	20	10	0	-10	-20	-30	-40	-50	ppm	PC 1.40









2c 1-5k #1-16 RT: 0.04-0.69 AV: 16 NL: 1.70E7 T: FTMS - p ESI Full ms [1000.00-5000.00]





### *in situ* <sup>11</sup>**B** NMR



										Current Data Parameters NAME 0203 EXPNO 41 PROCNO 1 F2 - Acquisition Parameters Date_ 20160203
										$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$
										======   CHANNEL f1     SF01   128.3776052 MHz     NUC1   11B     P1   10.00 usec     PLW1   52.00000000 W
 										F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB 0 PC 140
 30	20	10	0	-10	-20	-30	-40	-50	ppm	rC 1.40





<sup>1</sup>H NMR






<sup>11</sup>B NMR



						/9.61					
n	an far	Ymr47(14444-470-47	4-4wardarra	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	norme-magnet		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	Lawrench Margare Jacob	~~yuga~~~uurof~~uurof	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	$\begin{array}{c} \mbox{Current Data Parameters} \\ NAME & G1 2ME 0204 0202 (MeOD) \\ EXPNO & 251 \\ PROCNO & 1 \\ \hline \\ F2 - Acquisition Parameters \\ Date_ & 20160208 \\ Time & 17.02 \\ INSTRUM & av400 \\ PROBHD 5 mm PABBO BB/ \\ PULPROG & zg \\ TD & 5096 \\ SOLVENT & MeOD \\ NS & 1024 \\ DS & 0 \\ SWH & 51020.406 \ Hz \\ FIDRES & 10.011854 \ Hz \\ AQ & 0.0499408 \ sec \\ RG & 189.85 \\ DW & 9.800 \ usec \\ DE & 6.50 \ usec \\ DE & 6.50 \ usec \\ TE & 299.0 \ K \\ D1 & 0.05000000 \ sec \\ TD0 & 1 \\ \end{array}$
											======   CHANNEL f1   =======     SF01   128.3776052 MHz   NUC1   11B     P1   10.00 usec   PLW1   52.00000000 W
	20		10		10	20	20	40			F2 - Processing parameters     SI   32768     SF   128.3776161 MHz     WDW   EM     SSB   0     LB   10.00 Hz     GB   0     PC   1.40









0-1----

# Q Exactive High-Res Mass Spec







*in situ* <sup>11</sup>**B** NMR









<sup>1</sup>H NMR



	5.50 3.57 3.56 3.57 3.57 3.57 3.57 3.57 3.57 3.57 3.57	
		Current Data Parameters NAME G1 Glycerol 0204 0202 (MeOD) EXPNO 90 PROCNO 1 F2 - Acquisition Parameters Date_ 20160205 Time 12.37 INSTRUM av400 PROBHD 5 mm PABBO BB/ PULPROG zg30 TD 52882 SOLVENT MeOD NS 32 DS 0 SWH 8012.820 Hz FIDRES 0.151523 Hz AQ 3.2998369 sec RG 155.85 DW 62.400 usec DE 6.50 usec TE 299.0 K D1 5.00000000 sec TD0 1
9 8 7	1 2 5 4 5 6 3 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 1 2 3 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 2 3 1 1 1 1	<b>ppm</b> PC 1.00











2e #1-10 RT: 0.01-0.09 AV: 10 NL: 1.77E8 T: FTMS - p ESI Full ms [400.00-6000.00]





*in situ* <sup>11</sup>B NMR









<sup>1</sup>H NMR











1717.55

1800

1589.21

1400

1588.57

1600

G1 CA

100-

\*

0-

1000

1200

# Waters Mass Spec

2f 177 (5.967) 2: TOF MS ES-1632.55 1632.06 3198.91 3196.92 1631.58 -1633.06 1675.02 3281.19 3282.17 1675.57 1589.60 3283.20 3194.97 /1676.02 3284.19 1716.51

2600

2800

3000

3200

2200

2400

2000

3800

m/z

3285.11

3365.47

3400

3600

64.3



Waters Mass Spec

G1 CA









<sup>1</sup>H NMR



		7.5.51 7.3.63 7.3.59 7.3.59 7.3.59	3.57	L 3.34 L 3.34 L 3.34 L 3.09 L 3.09 L 3.06		Current Data Parameters NAME G1 PEG350 8 mg 0713 0710 (MeOD) EXPNO 32 PROCNO 1 F2 - Acquisition Parameters Date_ 20160713 Time 15.01 INSTRUM av400 PROBHD 5 mm PABBO BB/ PULPROG zg30 TD 52882 SOLVENT MeOD NS 32 DS 0 SWH 8012.820 Hz FIDRES 0.151523 Hz AQ 3.2998369 sec RG 155.85 DW 62.400 usec DE 6.50 usec TE 299.0 K
  9	 7	6 5 0 <sup>75</sup>	4 3 3257 4 3 35.5 23.0 23.0	2 1	 ppm	TD0 1   ======= CHANNEL f1 ======   SF01 400.1324008 MHz   NUC1 1H   P1 15.00 usec   PLW1 13.0000000 W   F2 - Processing parameters   SI 65536   SF 400.1300075 MHz   WDW EM   SSB 0   LB 0.30 Hz   GB 0   PC 1.00



30

20

10

0

-10

<sup>11</sup>B NMR

-15.74



Current Data Parameters NAME G1 PEG350 8 mg 0713 0710 (MeOD) EXPNO 30 PROCNO 1 F2 - Acquisition Parameters Date\_ 20160713 14.52 Time INSTRUM av400 PROBHD 5 mm PABBO BB/ zg 5096 PULPROG TD SOLVENT MeOD NS 1024 DS 0 51020.406 Hz SWH 10.011854 Hz 0.0499408 sec FIDRES AQ 189.85 RĜ 9.800 usec DW DE 6.50 usec TE 299.0 K D1 0.05000000 sec TD0 1 ====== CHANNEL f1 ======= SFO1 128.3776052 MHz 11B NUC1 P1 10.00 usec PLW1 52.0000000 W F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB PC 0 1.40 ..... т -20 -30 -40 -50 ppm









2i #1-16 RT: 0.01-0.14 AV: 16 NL: 2.75E6 T: FTMS - p ESI Full ms [400.00-6000.00]













*in situ* <sup>11</sup>**B** NMR









<sup>1</sup>H NMR







\* These peaks correspond to small impurities - boric acid and borates.





*in situ* <sup>11</sup>**B** NMR



											Current Data ParametersNAME0209EXPNO101PROCNO1F2 - Acquisition ParametersDate_20160209Time19.36INSTRUMav400PROBHD 5 mm PABBO BB/PULPROG $zg$ TD5096SOLVENTNoneNS1024DS0SWH51020.406 HzFIDRES10.011854 HzAQ0.0499408 secRG189.85DW9.800 usecDE6.50 usecTE299.0 KD10.05000000 secTD01
											=====   CHANNEL f1   ======     SFO1   128.3776052 MHz   NUC1   11B     P1   10.00 usec   PLW1   52.00000000 W
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~~~~							~~~~~		F2 - Processing parameters     SI   32768     SF   128.3776161 MHz     WDW   EM     SSB   0     LB   10.00 Hz     GB   0
	30	20	10	0	-10	-20	-30	-40	-50	ppm	rC 1.40






\* These peaks correspond to small impurities - boric acid and borates.





#### *in situ* <sup>11</sup>B NMR







Split peaks are due to the restricted rotational conformations of the molecule.9



<sup>1</sup>H NMR









Broad, split peaks are due to the restricted rotational conformations of the molecule.9





4800



21#1-57 RT: 0.01-0.49 AV: 57 NL: 2.33E7 T: FTMS - p ESI Full ms [900.00-5000.00]











Small impurities are present due to the commercial 1-hexanethiol used (95% pure).



F_	F ∕—S		2- 2[Et₃N	NH]+		<sup>11</sup> B { <sup>1</sup> H} NMR							
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	F	/1	2										BRUKER
								I					Current Data Parameters NAME Jan26-2016 EXPNO 40 PROCNO 1
													F2 - Acquisition Parameters   Date20160126   Time 12.53   INSTRUM av400   PROBHD 5 mm PABBO BB/   PULPROG zgdc.js   TD 5096   SOLVENT CD3CN   NS 1024   DS 0   SWH 51020.406 Hz   FIDRES 10.011854 Hz   AQ 0.0499408 sec   RG 189.85   DW 9.800 usec   DE 6.50 usec   TE 299.1 K   D1 0.03000000 sec   D11 0.03000000 sec   TD0 1
									Emergence   CHANNEL fl     SF01   128.3776052 MHz     NUC1   11B     P1   10.00 usec     PLW1   52.00000000 W				
											======================================		
 			·····								~~~~~~~~~		F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB 0
 <b>60</b>	50	40	30	20	10	0	-10	-20	-30	-40	-50	ppm	PC 1.40



Small impurities are present due to the commercial 1-hexanethiol used (95% pure).











in situ <sup>11</sup>B NMR























in situ <sup>11</sup>B NMR











1 2- 2[Et<sub>3</sub>NH]<sup>+</sup>











3c #1-16 RT: 0.01-0.14 AV: 16 NL: 1.92E6 T: FTMS - p ESI Full ms [400.00-6000.00]





in situ<sup>11</sup>B NMR


















3d 1-5k #1-16 RT: 0.04-0.69 AV: 16 NL: 5.34E6 T: FTMS - p ESI Full ms [1000.00-5000.00]





*in situ* <sup>11</sup>B NMR







-20

-40

-60

-80

-100











<sup>19</sup>F NMR









4800



3e 0.4-6k #1-18 RT: 0.01-0.16 AV: 18 NL: 9.85E5 T: FTMS - p ESI Full ms [400.00-6000.00]





*in situ* <sup>11</sup>**B** NMR









EM





#### <sup>19</sup>F NMR







Waters Mass Spec

G2 CA





# Waters Mass Spec



















3g #1-152 RT: 0.01-1.31 AV: 152 NL: 1.97E6 T: FTMS - p ESI Full ms [1000.00-6000.00]





in situ <sup>11</sup>B NMR











\* This peak corresponds to a small boric aicd impurity.





Waters Mass Spec

G2 CAG 5 mg/mL, 4:1 H2O:MeCN





Waters Mass Spec

#### G2 CAG 5 mg/mL, 4:1 H2O:MeCN





in situ <sup>11</sup>B NMR











<sup>11</sup>B NMR



-15.32 Current Data Parameters NAME G2 PEG350 1217 1209 (MeOD) EXPNO 3 PROCNO 1 F2 - Acquisition Parameters Date\_\_\_\_\_20151220 Time\_\_\_\_\_13.15 INSTRUM av400 PROBHD 5 mm PABBO BB/ PULPROG zg TD 5096 SOLVENT MeOD NS 1024 DS 0 SWH 51020.406 Hz FIDRES 10.011854 Hz 0.0499408 sec AQ RĠ 189.85 9.800 usec DW DE 6.50 usec ΤĒ 299.0 K D1 0.0500000 sec mannen TD0 1 www. ====== CHANNEL f1 ======= SFO1 128.3776052 MHz NUC1 11**B** P1 10.00 usec 52.0000000 W PLW1 F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz GB 0 PC 1.40 • • . 30 20 10 -20 -30 -40 -50 0 -10 ppm









3i 2 #1-20 RT: 0.01-0.17 AV: 20 NL: 2.87E5 T: FTMS - p ESI Full ms [2000.00-6000.00]




*in situ* <sup>11</sup>**B** NMR















in situ <sup>11</sup>B NMR









8

7

5

4

6

3

2

<sup>1</sup>H NMR



7.64 7.54 7.49 7.20	5.66	4.55	3.92	3.80	3.79	3.74	3.73	3.71	3.69	3.68	3.63	3.59	3.57	3.55	3.54	3.53	3.52	3.50	3.46	3.45	3.44	3.35	3.19
	Ľ	`_ 		 \																	<u> </u>	<u> </u>	Ú.

Current Data Parameters NAME G2 PEG2000 0203 0125 MeOD EXPNO 12 PROCNO 1 F2 - Acquisition Parameters Date\_\_\_\_\_20160204 Time 17.21 INSTRUM av400 PROBHD 5 mm PABBO BB/ PULPROG zg30 TD 52882 SOLVENT MeOD NS 32 DS 0 SWH 8012.820 Hz FIDRES 0.151523 Hz 3.2998369 sec AQ RĠ 48.1 62.400 usec DW DE 6.50 usec ΤĒ 299.0 K D1 5.0000000 sec TD0 1 ====== CHANNEL f1 ======= SFO1 400.1324008 MHz NUC1 1HP1 15.00 usec 13.0000000 W PLW1 F2 - Processing parameters SI 65536 SF 400.1300074 MHz WDW EM SSB 0 LB 0.30 Hz GB 0 PC 1.00 ----1 ppm



\* This peak corresponds to a small boric acid impurity.





in situ <sup>11</sup>B NMR



										Current Data Parameters
										NAME 0303 EXPNO 91 PROCNO 1
										$\begin{array}{rrrr} F2 - Acquisition Parameters \\ Date_ 20160303 \\ Time 20.25 h \\ INSTRUM av400 \\ PROBHD Z108618_0656 ( \\ PULPROG zg \\ TD 5096 \\ SOLVENT None \\ NS 1024 \\ DS 0 \\ SWH 51020.406 Hz \\ FIDRES 20.023708 Hz \\ AQ 0.0499408 sec \\ RG 189.85 \\ DW 9.800 usec \\ DE 6.50 usec \\ TE 299.0 \ K \\ D1 0.0500000 sec \\ TD0 1 \\ SFO1 128.3776052 \ MHz \\ NUC1 11B \\ P1 10.00 usec \\ PLW1 52.0000000 \ W \\ \end{array}$
 										F2 - Processing parameters SI 32768 SF 128.3776161 MHz WDW EM SSB 0 LB 10.00 Hz
 30	20	10	••••• 0	-10	-20	-30	-40	-50	ppm	GB 0 PC 1.40



-80

-60

-20

-40

-100

-120

-140

24.0

23.8

-160



	Current Data Parameters NAME 0303 EXPNO 90 PROCNO 1
	F2 - Acquisition Parameters   Date_ 20160303   Time 20.20 h   INSTRUM av400   PROBHD Z108618_0656 (   PULPROG zgflqn30   TD 262144   SOLVENT None   NS 64   DS 0   SWH 150000.000 Hz   FIDRES 1.144409 Hz   AQ 0.8738133 sec   RG 189.85   DW 3.333 usec   DE 6.50 usec   TE 299.0 K   D1 2.00000000 sec   TD0 1   SFO1 376.4983660 MHz   NUC1 19F   P1 14.50 usec   PLW1 17.0000000 W
	F2 - Processing parameters SI 262144 SF 376.4983660 MHz WDW EM SSB 0
	LB 1.00 Hz GB 0 PC 1.00
ppm	











#### Q Exactive High-Res Mass Spec





#### Q Exactive High-Res Mass Spec

3I #1-16 RT: 0.01-0.14 AV: 16 NL: 1.10E5 T: FTMS - p ESI Full ms [400.00-6000.00] 1840.3178 z=3 100 95 1840.6522 90 85 80 75 z=3 1839.9836 z=3 70 65 60 55 50 45 40 1840.9840 z=3 Relative Abundance 1841.3176 z=3 1839.6510 z=3 1839.3185 z=3 35 30 25 1841.6507 1841.9860 -20 15 10 z=3 z=3 1838.9851 z=3 1842.3157 1838.6512 z=3 z=3 1843.3113 1842.6517 z=? 1837.3446 5 1836.3261 1838.3165 z=3 z=3 1837.0509 z=? 1837.6916 z=? z=? Z=? Μ 1838.0 1838.5 1843.0 1844.0 1836.5 1837.0 1837.5 1839.0 1839.5 1840.0 1840.5 1841.0 1841.5 1842.0 1842.5 1843.5

m/z





#### 2D DOSY <sup>1</sup>H NMR











#### 2D DOSY <sup>1</sup>H NMR







#### Stability studies of 2i under biologically relevant conditions

**Cell culture media/fetal bovine serum:** 14.8 mg of **2i** was dissolved in 500  $\mu$ L of Milli-Q water. 100  $\mu$ L of this solution was added to 500  $\mu$ L of serum media (440  $\mu$ L cell culture media and 60  $\mu$ L fetal bovine serum). This mixture was vortexed and then transferred to an NMR tube and monitored over 5 days at room temperature by <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. This sample was then incubated at 37 °C for an additional 5 days and subjected to analysis *via* <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. No significant change was observed by NMR spectroscopy.

**pH 5:** 14.8 mg of **2i** was dissolved in 500  $\mu$ L of Milli-Q water. 100  $\mu$ L of this solution was added to 500  $\mu$ L of a 0.1 M citric acid/sodium citrate buffer at pH 5.0. This mixture was vortexed and then transferred to an NMR tube and monitored over 5 days at room temperature by <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. NMR spectroscopy suggests that the structural integrity is maintained. We note that we observed small impurities corresponding to boric acid and borates by <sup>11</sup>B NMR spectroscopy as well as some peak broadening in <sup>11</sup>B and <sup>19</sup>F NMR spectra due to the oxidation of **2i** from the 2- to the 1- oxidation state over time.

**pH 7:** 14.8 mg of **2i** was dissolved in 500 μL of Milli-Q water. 100 μL of this solution was added to 500 μL of a 0.1 M Tris/HCl buffer at pH 7.0. This mixture was vortexed and then transferred to an NMR tube and monitored over 5 days at room temperature by <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. NMR spectroscopy suggests that the structural integrity is maintained. We note that we observed small impurities corresponding to boric acid and borates by <sup>11</sup>B NMR spectroscopy as well as some peak broadening in <sup>11</sup>B and <sup>19</sup>F NMR spectra due to the oxidation of **2i** from the 2- to the 1oxidation state over time.

**pH 9:** 14.8 mg of **2i** was dissolved in 500  $\mu$ L of Milli-Q water. 100  $\mu$ L of this solution was added to 500  $\mu$ L of a 0.1 M Tris/HCl buffer at pH 9.0. This mixture was vortexed and then transferred to

an NMR tube and monitored over 5 days at room temperature by <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. NMR spectroscopy suggests that the structural integrity is maintained. We note that we observed small impurities corresponding to borates by <sup>11</sup>B NMR spectroscopy as well as some peak broadening in <sup>11</sup>B and <sup>19</sup>F NMR spectra due to the oxidation of **2i** from the 2- to the 1- oxidation state over time.

**2-Mercaptoethanol:** 16.9 mg of **2i** was dissolve in 2.82 mL of D<sub>2</sub>O. 500  $\mu$ L of this solution was added to 100  $\mu$ L of a 120 mM 2-mercaptoethanol D<sub>2</sub>O solution. This mixture was vortexed and then transferred to an NMR tube and monitored over 11 days at room temperature by <sup>1</sup>H, <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. After 11 days, this sample was subjected to mass spectrometry analysis. Both NMR spectroscopy and mass spectrometry suggest that the structural integrity is maintained. We note that we observed a small boric acid impurity by <sup>11</sup>B NMR spectroscopy.

**Glutathione:** 16.9 mg of **2i** was dissolve in 2.82 mL of D<sub>2</sub>O. 500  $\mu$ L of this solution was added to 100  $\mu$ L of a 12 mM glutathione D<sub>2</sub>O solution. This mixture was vortexed and then transferred to an NMR tube and monitored over 11 days at room temperature by <sup>1</sup>H, <sup>11</sup>B and <sup>19</sup>F NMR spectroscopy. After 11 days, this sample was subjected to mass spectrometry analysis. Both NMR spectroscopy and mass spectrometry suggest that the structural integrity is maintained. We note that we observed a small boric acid impurity by <sup>11</sup>B NMR spectroscopy as well as some peak broadening in <sup>11</sup>B and <sup>19</sup>F NMR spectra due to the oxidation of **2i** from the 2- to the 1- oxidation state over time.







# Stability of 2i in Serum <sup>19</sup>F NMR









# Stability of 2i at pH 5 <sup>19</sup>F NMR









# Stability of 2i at pH 7 <sup>19</sup>F NMR









# Stability of 2i at pH 9 <sup>19</sup>F NMR







# Stability of 2i in 2-Mercaptoethanol <sup>1</sup>H NMR



BRÚKÉR








# Stability of 2i in 2-Mercaptoethanol - Day 11 Waters Mass Spec





# Stability of 2i in Glutathione <sup>1</sup>H NMR













## Stability of 2i in Glutathione - Day 11 Waters Mass Spec



### Plot of Conjugation Progress of Boc-cysteine (g) onto Clusters 2/3



### **Computational work**

#### **A. PEGylated OCNs**

PEGylated nanoparticles (NPs) 2i-k and 3i-k (see Table 2) were modeled using molecular dynamics (MD) simulations in: *i*) water with counter ions and *ii*) a buffer solution of HPO<sub>4</sub><sup>2-</sup> and  $H_2PO_4^-$  at a total 0.08 M concentration, where the ratio of the two ions was used matched pH 7.4. The MD simulations were performed with NAMD<sup>10</sup>, using the CHARMM force field<sup>11-16</sup>. Ab initio calculations were done with Gaussian0917 to determine unknown parameters for the dodecaborate cluster center and the non-PEGylated (2 or 3 type ligand) section of the ligand. The boron center was optimized using a HF/6-31g level of theory, with partial charges derived with a ChelpG algorithm<sup>18</sup>. Bonds, angles, and dihedrals force constants containing boron atoms were chosen to have relatively large values, approximately equal to those of double bonded or aromatic carbons, so that the boron center would be rigid. The type 2 and 3 ligands had their bond and angle parameters determined at the MP2/6-31g(d)//HF/6-31g level of theory with VMD Force Field Toolkit plugin<sup>19</sup>. Unknown dihedral parameters were chosen based on similar atom types in the CHARMM force field<sup>11-16</sup>. Partial charges were determined through the ChelpG aglorithm<sup>18</sup>. Amide and PEGylated geometries, parameters, and charges were taken from the CHARMM force field<sup>11–16</sup>.

Each of the 6 NPs was separately simulated in water and ionic solutions. Each system is first minimized for 10,000 steps. Afterwards it is heated to 310 K, with 1 K increments per 20 steps until the system reaches a temperature of 310 K, when a pre-equilibration is done. Simulations are performed in an NPT ensemble, at 310 K and a pressure of 1 atm, with Langevin dynamics and a damping constant of 0.01 ps<sup>-1</sup>. Langevin piston is used with a period of 200 fs and decay of 100 fs. Particle Mesh Ewald<sup>20</sup> is used for long range electrostatic interactions with a grid

spacing of 1.0. Short range interactions are performed with the 12-6 Lennard-Jones potential, using a switching function. Velocity Verlet integration is used with the SHAKE algorithm and a time step of 2 fs. Data and snapshots are recorded every 10 ps or 5,000 steps. Simulation times of 25 ns for the water solution and 30 ns for the salt system were used, respectively. Supplementary Figures 1 and 2 show snapshots of PEGylated NPs in water (21 ns) and in the ionic solution (31 ns), respectively. Notice that as the chain length increases, the chains are fluctuating significantly to the extent that the chain distributions become asymmetric. In the following, we describe some characteristics of these systems.



Supplementary Figure 1. Nanoparticles snapshots in water after 21 ns of simulations. Scale bar is 1 nm. A) 2i B) 3i C) 2j D) 3j E) 2k F) 3k.



Supplementary Figure 2. Nanoparticles snapshots in 0.08 M buffer solution at pH=7.4 (salt) after
31 ns of simulations. Scale bar is 1 nm. A) 2i B) 3i C) 2j D) 3j E) 2k F) 3k.

We use the simulated trajectories of the NPs to calculate the radial distribution functions (RDF), g(r), from Eqn. 1. It gives the relative probability of finding the j<sup>th</sup> atom at a distance r from the i<sup>th</sup> atom with respect to the bulk density:

$$g(r) = \frac{1}{V\rho_N} \sum \delta(r - r_{ij}) . \tag{1}$$

In Eqn. 1,  $\delta$  is a delta function,  $r_{ij}$  is the distance of i<sup>th</sup> and j<sup>th</sup> atoms, and V is a volume,  $\int 4\pi r^2 dr$ , used in a normalization, and  $\rho_N$  is the number density of the used species (the number of atoms N<sub>0</sub> used in Eqn. 1 divided by the volume of the simulation box). We use Eqn. 1 when we analyze the distribution of C terminal atoms, which are fixed for a given number of ligands (12). When, we consider the distribution of all PEG-chain oxygens (varying number), we remove N<sub>0</sub> (equal the total number of PEG chain oxygens) from  $\rho_N$ , by multiplying Eqn.1 by N<sub>0</sub>, to get g'(r), where we account for the growing distributions for longer PEG chains (more oxygens; system volume is fixed).



**Supplementary Figure 3.** RDFs of **2i–k** and **3i–k** NPs. g'(r) calculated for A) boron-PEG oxygen atoms in water and B) boron-PEG oxygen atoms in ionic solution. g(r) calculated for C) boron-terminal C atoms in water and D) boron-terminal C atoms in ionic solutions.

In Supplementary Figure 3, we have calculated g'(r) for (A, B) all the oxygens in PEGylated chains and g(r) for (C, D) terminal carbon atoms of the PEGylated chains. All the cases were calculated with respect to all the boron atoms. We can clearly see that as the chain becomes longer, the oxygen (A, B) distributions become wider and their peaks,  $r_{max}$ , become slightly shifted to higher values. Steric effects prevent longer PEGylated chains from folding and wrapping close to the B core, therefore, preventing them from significantly affecting  $r_{max}$ . The systems present in

water and ionic solutions have almost the same PEG-oxygens distributions. On the contrary, in the terminal carbon (CD) distributions, the peaks maxima,  $r_{max}$ , are significantly shifted to higher values with the chain lengths, since the terminal C atoms are further away from the NPs cores, which they cannot reach. In these distributions, we can also see some differences between water and ionic solution cases, revealing that the terminal atoms in long PEGylated chains are slightly more outstretched in ionic solutions.

The g'(r) distributions (Supplementary Figure 3 A, B) are similar for the **2** and **3** types of ligands, except of some deformations present in the **3** types. These deformations slightly shift the **3** type peaks ( $r_{max}$ ) to smaller values. For all but **2k** and **3k** terminal carbon RDFs, **3** type ligands have consistently smaller  $r_{max}$  values than their **2** type counterparts (Supplementary Figure 3), even though **3** has an extra aromatic group, slightly increasing the maximum possible ligand length. The extra aromatic group in **3** ligands enhances  $\pi$ - $\pi$  stacking interactions between the ligands, thus causing the net length to decrease. The split peak in **2i** could be related to the fact that the B shell front and back sides can contribute by separate peaks.

The hydrodynamic radii of the studied NPs were estimated from the regions of decaying g'(r) (half value compared to r<sub>max</sub>) for the cases (A–B) (all oxygens). In water, the hydrodynamic radii of **2i** and **3i** are 12 Å; **2j** and **3j** are 15 Å; **2k** and **3k** are 20 Å. In the ionic solution, **2i**, **3i**, **2j**, and **3j** have very similar sizes as in water. At certain times, there are some chains on **2k** or **3k** that extend outwards, but most of the other chains are folded (Supplementary Figures 1 E, F and 2 E, F). Interestingly, the maxima of distributions for the terminal C atoms in Supplementary Figure 3 C, D match relatively well to the hydrodynamic radii. One can assume that the terminal C atoms are distributed at the surface of the NPs, revealing thus their radii.

To confirm the previous results, next, the radii of gyration,  $\langle r_{gyr} \rangle$ , are also calculated for NPs using Eqn. 2:

$$r_{gyr} = \sqrt{\frac{I}{m}} = \sqrt{\frac{\sum_{i=atoms} m_i (\vec{r}_i - \vec{r}_{com})^2}{\sum_{i=atoms} m_i}} \quad . \tag{2}$$

Here, I is the moment of inertia of the molecule, m is the total mass of the molecule formed by individual contributions,  $m_i$ , of atoms shifted with respect to a molecular center of mass,  $r_i$ - $r_{com}$ . Time averaged  $\langle r_{gyr} \rangle$  was calculated by using equation 2 every 10 ps over 26 ns trajectory (water) or 34 ns trajectory (salt solution) and then averaged. Standard deviations and confidence intervals were also computed.

**Supplementary Table 2.** Radii of gyration, <r<sub>gyr</sub>>, and their confidence intervals for PEGylated species in water and salt solutions.

Molecule	Solvent	$< r_{gyr} > (Å)$
2i	water	$11.5\pm0.9$
2j	water	$15.0\pm1.7$
2k	water	$20.7\pm2.2$
3i	water	$12.1 \pm 1.2$
3j	water	$14.7 \pm 1.3$
3k	water	$21.1\pm2.0$
2i	ionic solution	$11.7 \pm 1.0$
2j	ionic solution	$14.7\pm2.0$
2k	ionic solution	$21.0\pm2.5$
3i	ionic solution	$12.2 \pm 1.5$
<u> </u>	ionic solution	$14.8 \pm 1.6$
3k	ionic solution	$22.1\pm4.5$

Supplementary Table 2 shows the radii of gyration,  $\langle r_{gyr} \rangle$  and their >99.5 % confidence intervals for PEGylated species in water and salt solutions. As expected, **2i** and **3i** molecules have the smallest diameters, whereas **2k** and **3k** have the largest diameters in both environments. **2i** and **3i**  molecules, with 7 PEGylated oxygens per ligand have diameters of more than 2 nm; 2j and 3j, with 16 PEGylated oxygens per ligand, less than 3 nm; 2k and 3k, with 43 oxygens per ligand, more than 4 nm. NPs with the type 3 ligands tend to have a slightly larger diameter than those with the type 2 ligands. This size increase could be due to the extra aromatic group in type 3 ligands, which is absent in the 2 type ligands.  $\langle r_{gyr} \rangle$  does not change appreciably between the two environments. However, 2k and 3k ligands are slightly more outstretched in the ionic solutions.



Supplementary Figure 4. Distributions of  $r_{gyr}$  in a) water and b) ionic solutions.

Supplementary Figures 4a and b show the distributions of  $r_{gyr}$  in water and salt solutions, respectively. The distributions are asymmetrically broadened at higher values for all molecules, especially for long chains. This reflects that a few chains could extend and then fold back. Comparing the radii of gyration,  $\langle r_{gyr} \rangle$ , from Supplementary Table 2 and Supplementary Figure 4 with the above hydrodynamic radii and the most likely positions of terminal C atoms, we can see that all these parameters are in good agreement.

#### **B.** Sugar-coated nanoparticles – protein binding

MD simulations were also performed to investigate multivalent binding of sugar-coated nanoparticles and proteins. Concanavalin A (Con A) was chosen as the target protein to bind with

multivalent sugar-coated particles (SP) and monovalent β-D glucose (G), respectively. Con A forms quaternary structures, giving at pH 7 a tetramer, having four carbohydrates binding sites (hydrogen bonds)<sup>21</sup>. In each Con A, up to 15 amino acids can be involved in the carbohydrate binding, while for the monosaccharide binding only five amino acids are involved, including Asn 14, Leu 99, Tyr 100, Asp 208, Arg 228<sup>22</sup>. In our simulations, the tetramer structure of Con A used was based on X-ray diffraction data (PDB code 1ONA)<sup>22</sup>. Supplementary Figures 5A and B show the structures of tetramer of Con A with SPs and β-D glucose after 20 ns simulation. The metals manganese (magenta ball) and calcium (cyan ball) were added in Con A according to its metal binding sites<sup>22</sup>. The monosaccharide binding sites are distinguished from the backbone of Con A by different colors (shown in Supplementary Figure 5). The Con A tetramer has four binding positions. We name the top right position as binding position 1 (B1), bottom right as B2, top left as B3, and bottom left as B4.



Supplementary Figure 5. A) Tetramer of Con A and sugar-coated particles. B) Tetramer of Con A and  $\beta$ -D glucoses. Details of glucose binding shown in both cases.

For the NPs binding, three SPs (SP1, SP2 and SP4) were initially put near the binding sites of chosen monomers. The last SP (SP3) was placed in the cavity between the B1 and B3 binding positions. For the  $\beta$ -D glucose binding, three glucose molecules (G1, G2 and G3) are separately placed at the binding B1, B2 and B3 positions, while the last glucose molecule (G4) was placed between the B3 and B4 binding position. The two systems were immersed in water together with the counter-ions and the simulations were performed with NAMD<sup>10</sup>.

The bond, angle and dihedral parameters of protein, SPs (nanoparticle **2l** in Table 2) and  $\beta$ -D glucose were implemented from the CHARMM<sup>11–16</sup> force field. The parameters for the boron core and ligands were used the same as in the PEGylated calculations. The nonbonding parameters of Mn<sup>2+</sup> ions were based on the calculations of Babu *et al.*<sup>23</sup>. Nonbonding interactions of SPs were calculated using a cut-off distance of 10 Å, whereas long-range electrostatic interactions were calculated by the PME method<sup>20</sup> in the presence of periodic boundary conditions. The systems were simulated in the NPT ensemble, using a Langevin dynamics with a damping constant of 0.01 ps<sup>-1</sup> and a time step of 1 fs.

First, we modeled the coupling between SPs and the Con A tetramer. At each simulation time, we have calculated a distance between each sugar binding site and its nearest ligand in the SP. Supplementary Figure 6 shows a time-dependent distance between the nearest SPs ligand and the Con A tetramer. During the 20 ns simulations, SP1 and SP2 have an average distance of 4 Å, while SP3 and SP4 have an average distance of about 10 Å. Because the initial position of SP3 is far from any binding site, it can't bind during the short simulations. From Supplementary Figure 5A, we can see that SP3 competes with SP1 for the B1 position, while SP4 shows a different trend. Within 1 ns, SP4 comes near to the Con A tetramer and binds to it. Then, it leaves away and binds



again at 4 ns, when the binding lasts for about 4 ns. After 12 ns, SP4 binds to the Con A tetramer again.



Supplementary Figure 6 reveals that when SPs bind to the Con A tetramer their binding distance is about 1.8-2 Å. SPs occasionally gain and preserve for significant time periods these small binding distances. Supplementary Figure 5 A(a-c) show details of SP1, SP2 and SP4 binding to their binding sites. We can see that in all the cases only one of the SPs ligands binds to the nearby binding site, composed of Asn 14, Leu 99, Tyr 100, Asp 208, Arg 228, which is the monosaccharide binding site shown in different color in Supplementary Figure 5 A(a-c). Therefore, there is always one ligand of SPs which performs like a monosaccharide when binding to the Con A tetramer. Because the SPs have several ligands, when one ligand leaves, another



nearby ligand comes and binds, which increases the binding probability of SPs. In this way, SPs act like multivalent binders.

**Supplementary Figure 7.** Nearest distances between  $\beta$ -D glucose molecules and the Con A tetramer.

In order to compare the binding ability of SPs and  $\beta$ -D glucose systems, we simulated binding of  $\beta$ -D glucose and the Con A tetramer. Supplementary Figure 7 shows the nearest distance between  $\beta$ -D glucose and Con A as a function of time. Supplementary Figure 7 shows that G1 only binds to Con A at the first 1 ns and then leaves. G2 only binds at the very beginning and it doesn't bind later; G3 binds to Con A for about 4 ns at the beginning and after that it leaves away; G4 shows weak binding during the first 4 ns. The average distance between all the  $\beta$ -D glucose molecules and the Con A tetramer is more than 20 Å, except G3 whose average distance is about 12 Å. Supplementary Figure 5B(a-c) shows details of  $\beta$ -D glucose and the Con A tetramer binding. When  $\beta$ -D glucose binds to Con A, it binds to the typical monosaccharide binding sites. Because  $\beta$ -D glucose is monovalent, when one  $\beta$ -D glucose leaves, another  $\beta$ -D glucose from the surrounding solution might come nearby and bind. Overall, monovalent  $\beta$ -D glucose molecules show shorter binding times and longer binding distances than SPs.

### References

- Safina, G., Duran, I. B., Alasel, M. & Danielsson, B. Surface plasmon resonance for real-time study of lectin–carbohydrate interactions for the differentiation and identification of glycoproteins. *Talanta* 84, 1284–1290 (2011).
- Sheldrick, G. M. SHELXT integrated space-group and crystal-structure determination. *Acta Cryst.* A71, 3–8 (2015).
- 3. Sheldrick, G. M. A short history of SHELX. Acta Cryst. A64, 112–122 (2008).
- Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2 : a complete structure solution, refinement and analysis program. *J. Appl. Cryst.* 42, 339–341 (2009).
- Wixtrom, A. I. *et al.* Rapid synthesis of redox-active dodecaborane B<sub>12</sub>(OR)<sub>12</sub> clusters under ambient conditions. *Inorg. Chem. Front.* 3, 711–717 (2016).
- 6. Messina, M. S. *et al.* Visible-light-induced olefin activation using 3D aromatic boron-rich cluster photooxidants. *J. Am. Chem. Soc.* **138**, 6952–6955 (2016).
- Becer, C. R. *et al.* Clicking pentafluorostyrene copolymers: synthesis, nanoprecipitation, and glycosylation. *Macromolecules* 42, 2387 – 2394 (2009).
- Hood, C. A. *et al.* Fast conventional Fmoc solid-phase peptide synthesis with HCTU. *J. Pept. Sci.* 14, 97-101 (2008).
- 9. Zou, Y. *et al.* Convergent diversity-oriented side-chain macrocyclization scan for unprotected polypeptides. *Org. Biomol. Chem.* **12**, 566–573 (2014).
- 10. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

- 11. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–616 (1998).
- 12. Vanommeslaeghe, K. & MacKerell, A. D. Automation of the CHARMM general force field (CGenFF) I: bond perception and atom typing. *J. Chem. Inf. Model.* **52**, 3144–3154 (2012).
- Vanommeslaeghe, K., Raman, E. P. & MacKerell, A. D. Automation of the CHARMM general force field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* 52, 3155–3168 (2012).
- Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* 33, 2451–2468 (2012).
- 15. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* 25, 1400– 1415 (2004).
- Vanommeslaeghe, K. *et al.* CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 31, 671–690 (2010).
- 17. Gaussian 09, Revision D.01, Frisch, M. J. et al. Gaussian, Inc., Wallingford, CT (2013).
- Breneman, C. M. & Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11**, 361–373 (1990).
- 19. Mayne, C. G., Saam, J., Schulten, K., Tajkhorshid, E. & Gumbart, J. C. Rapid parameterization of small molecules using the force field toolkit. *J. Comput. Chem.* **34**, 2757–2770 (2013).

- 20. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
- Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. Legume lectin structure. *Biochim. Biophys.* Acta - Protein Struct. Mol. Enzymol. 1383, 9–36 (1998).
- 22. Loris, R., Maes, D., Poortmans, F., Wyns, L. & Bouckaert, J. A structure of the complex between Concanavalin A and methyl-3,6-di-O-(α-D-mannopyranosyl)-α-D-mannopyranoside reveals two binding modes. *J. Biol. Chem.* **271**, 30614–30618 (1996).
- 23. Babu, C. S. & Lim, C. Empirical force fields for biologically active divalent metal cations in water. *J. Phys. Chem. A* **110**, 691–699 (2006).